Surrogate Modeling and Global Sensitivity Analysis towards Efficient Simulation of Nuclear Reactor Stochastic Dynamics

by

Gregory Allen Banyay

Bachelor of Science in Mechanical Engineering, Ohio University, 2004

Master of Science in Mechanical Engineering, Ohio University, 2006

Submitted to the Graduate Faculty of

Swanson School of Engineering in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2019

UNIVERSITY OF PITTSBURGH

SWANSON SCHOOL OF ENGINEERING

This dissertation was presented

by

Gregory Allen Banyay

It was defended on

February 8, 2019

and approved by

Jeen-Shang Lin, Ph.D.
Associate Professor, Department of Civil & Environmental Engineering

Vikas Khanna, Ph.D.
Associate Professor, Department of Civil & Environmental Engineering

Michael D. Shields, Ph.D.
Assistant Professor, Department of Civil Engineering (Johns Hopkins University)

Dissertation Director: John C. Brigham, Ph.D.
Associate Professor, Department of Civil & Environmental Engineering
Associate Professor, Department of Engineering (Durham University)

Surrogate Modeling and Global Sensitivity Analysis towards Efficient Simulation of Nuclear Reactor Stochastic Dynamics

Gregory Allen Banyay, Ph.D.

University of Pittsburgh, 2019

Surrogate modeling is used to support global sensitivity analyses (GSA) for the modeling and simulation of nuclear reactor assembly structural dynamics to demonstrate the pertinence of such methods to this application as well as the significant physical insights provided by GSA. In addition to the knowledge gained related to the system sensitivity, insight gained from the accuracy of the GSA results may be used to compare with goodness-of-fit metrics that are traditionally used for verification of the surrogate model. The coupled use of surrogate modeling and GSA reduces the number of full-order simulations required, substantially reducing total computational cost. This work focuses on the use of Gaussian Process surrogates in particular, and examines the robustness of these techniques to evaluate sensitivity by considering a variety of design of experiment strategies used to create the surrogate models.

Numerical experiments based upon two finite element models representing stochastic dynamics for a pressurized water reactor, are used to evaluate the relationship between sensitivities computed from a full-order model versus those computed from a surrogate model, highlighting the effectiveness of utilizing GSA and surrogate modeling. For the examples presented herein the historical significance of both forcing function characterization and model parameter definition is substantiated, in terms of the GSA providing insight as to dominant contributors to structural dynamic behavior. For large sample sizes, negligible variation in the resultant sensitivities is shown with respect to the particular method by which a computational design of experiment is constructed to train the surrogates, that demonstrates stability and veracity of the results. For small sample sizes, the use of Latinized Partially Stratified Sampling (LPSS) provided surrogates and associated sensitivities with lower error as compared to Latin Hypercube Sampling (LHS) and sampling via the Fourier Amplitude Sensitivity Test (FAST). Differences in GSA results imparted by examining time-domain versus spectral acceleration results, as

well as increasing model parameter variation further illustrated the effectiveness of advanced sampling methods. Furthermore, the use of adaptive sampling and aggregate surrogate modeling techniques are introduced, with which incremental improvements were realized regarding the number of samples required to achieve accurate surrogate models.

Table of Contents

Nomenclature ..... **xiv**

Preface ..... **xv**

1.0 Introduction **1**

1.1 Dissertation Outline and Research Questions Addressed **13**

2.0 Methodology **15**

2.1 Random Vibration **17**

2.1.1 Stationary Random Vibration **17**

2.1.2 Non-Stationary Random Vibration **21**

2.1.2.1 Finite Element Formulation **21**

2.1.2.2 Acceleration Response Spectra **22**

2.2 Global Sensitivity Analysis **26**

2.2.1 Fourier Amplitude Sensitivity Test (FAST) Method **27**

2.2.2 Correlation based Methods **29**

2.3 Surrogate Modeling **31**

2.3.1 Gaussian Process Regression **31**

2.3.2 Genetic Aggregation **33**

2.3.3 Verification Criteria **37**

2.4 Design of Experiment Sampling Strategies **39**

2.4.1 Latin Hypercube Sampling **40**

2.4.2 Latinized Partially Stratified Sampling **41**

2.4.3 Adaptive Sampling **43**

2.5 Integrated Computational Framework **48**

3.0 Application to Stationary Random Vibration for Upper Internals Assembly 50
3.1 Finite Element Analysis Description 52
3.2 Sensitivity Analysis of Stationary Random Vibration for UI Assembly 56
3.2.1 Sensitivity Analysis of Full-Order Model using FAST 56
3.2.2 Surrogate-based Sensitivity Analysis 58
3.2.2.1 Surrogate trained by FAST Samples 59
3.2.2.2 Surrogate trained by LHS and LPSS 61
3.3 Discussion from Evaluation of UI Flow-Induced Vibration 67
3.3.1 Surrogate Model Accuracy 67
3.3.2 GSA Comparison from Full-Order Model versus Surrogate Models.... 69
3.3.3 Conclusions from Upper Internals FIV Evaluation 71
4.0 Application to Non-Stationary Random Vibration of Reactor System 73
4.1 Finite Element Analysis Description 74
4.2 Sensitivity Analysis of Non-Stationary Random Vibration for SFEM..... 82
4.2.1 Sensitivity Analysis of Full-Order Model using FAST 85
4.2.2 Surrogate-based Sensitivity Analysis 86
4.2.3 Comparison Studies 93
4.2.3.1 Effect of Increasing Input Parameter Variance 94
4.2.3.2 Evaluation of Acceleration Response Spectra Variation..... 103
4.2.4 Surrogate Model Accuracy 112
4.3 Conclusions from Evaluation of System Finite Element Model 115
5.0 Effectiveness of Surrogate Modeling Methodological Changes..... 119
5.1 Effect of Aggregate Surrogate Model 120
5.1.1 Correlation of Cross-Validation and Verification Points..... 124
5.2 Effect of Adaptive Sampling 126

5.3 Correlation versus Variance-Based Global Sensitivity Methods 131

6.0 Conclusions of Dissertation 137

Matlab Code for Latinized Partially Stratified Sampling..... 142

Discussion of Spectral Response Method..... 146

Bibliography 152

List of Tables

Table 1.1 Cost Benefit Analysis Variables 10

Table 1.3 Outline for Dissertation 13

Table 1.4 Research Questions 14

Table 2.1 Time Integration Coefficients used for Transient Finite Element Method 22

Table 2.2 Surrogate Model Goodness of Fit Criteria 38

Table 2.3 LSS Method Procedure 42

Table 3.1 Forcing Functions Applied to Upper Internals Finite Element Model 56

Table 3.2 Setup of Partially Stratified Sampling Design 61

Table 3.3 Sensitivity Indices Computed from Surrogate Models 63

Table 4.1 SFEM Parameters 77

Table 4.2 Analysis of Variance Variability Table Results 78

Table 4.3 Subdomains for LPSS of SFEM 79

Table 4.4 Factorial DOE Result for Pressure Vessel Acceleration with ±25% Input Parameter Variation 102

Table 5.1 Comparison of Surrogate Error Metrics between Kriging and Aggregate Surrogate Model using LPSS 121

Table 5.2 Comparison of Surrogate Error Metrics between Kriging and Aggregate Surrogate Model using LHS 121

Table 5.3 Genetic Aggregation Training Generations for output parameter P17 123

Table 5.4 Genetic Aggregation Training Generations for output parameter P18 123

Table 5.5 Cases Considered to Study the Effect of Adaptive Sampling 127

Table 5.6 Adaptive Refinement Point Placement 131

List of Figures

Figure 1.1 Influence Diagram and Inputs for Postulated Cost-Benefit Analysis 9

Figure 1.2 Illustration of Cost-Benefit Analysis between Traditional Approach and the use of Sensitivity Analysis 10

Figure 2.1 Flow Chart of Analytical Workflow involving model definition, Computational DOE, Surrogate Training, and GSA 16

Figure 2.2 Flow Chart of Genetic Aggregation Method 36

Figure 2.3 Illustration of Adaptive Sampling Principle 44

Figure 2.4 Integrated Workflows 49

Figure 3.1 Upper Internals Assembly for the AP1000 Plant (from (Westinghouse Electric Company, 2011) 51

Figure 3.2 Upper Internals Model, Mesh, and Boundary Conditions 53

Figure 3.3 Forcing Function Non-Dimensional Power Spectral Density (with beam forced response data overlaid) 55

Figure 3.4 Full-order Model GSA Convergence 57

Figure 3.5 GSA from Kriging Surrogates trained with FAST at varying Sample Sizes 60

Figure 3.6 FAST-trained Surrogate-based Sensitivities vs Benchmark 60

Figure 3.7 Comparison of Sensitivities from LHS-Trained Surrogate Models 64

Figure 3.8 Comparison of Sensitivities from LPSS-Trained Surrogate Models 64

Figure 3.9 Comparison of LHS-trained and LPSS-trained Surrogate-based Sensitivities to Benchmark for Base Motion loads 65

Figure 3.10 Comparison of LHS-Trained and LPSS-Trained Surrogate-based Sensitivities to Benchmark for Cross-Flow Loads 65

Figure 3.11 Surrogate-based Sensitivities vs. Benchmark for Base Motion Loads 66

Figure 3.12 Surrogate-based Sensitivities vs. Benchmark for Cross-Flow Loads 66

Figure 3.13 Surrogate Model Verification by Comparison to Test Set using Root Mean Squared Error (Strain) 68

Figure 4.1 Illustration of Typical Pressurized Water Reactor Assembly 74

Figure 4.2 Illustration of System Finite Element Model 75

Figure 4.3 Down-Selection of Parameters for SFEM Sensitivity Analysis 80

Figure 4.4 Identification of Input Parameter Interactions for SFEM 81

Figure 4.5 Single Factor Sensitivity of 6 Outputs to Input P6 83

Figure 4.6 Single Factor Sensitivity of 6 Outputs to Input P8 84

Figure 4.7 Single Factor Sensitivity of 6 Outputs to Input P15 84

Figure 4.8 Full-Order Model GSA Results for SFEM using FAST 86

Figure 4.9 Surrogate-based GSA Results for SFEM, Compared to Full-Order GSA Results using LHS and LPSS, Respectively 87

Figure 4.10 Difference between Kriging-based and Full-Order GSA Results of SFEM (Bar Graph) using LHS and LPSS, Respectively 89

Figure 4.11 Difference between Kriging-based and Full-Order GSA Results of SFEM (Line Graph) 90

Figure 4.12 Illustration of Kriging Surface for Output P17 with respect to $\boldsymbol{N6}$ (P12) and $\boldsymbol{k9}$ (P15) 91

Figure 4.13 Illustration of Kriging Surface for Output P17 with respect to $\boldsymbol{k10}$ (P16) and $\boldsymbol{k9}$ (P15) 91

Figure 4.14 Slices through Kriging Surface for Output P17 with respect to $\boldsymbol{k10}$ (P16) and $\boldsymbol{k9}$ (P15) 92

Figure 4.15 Interaction Plot for Lateral Acceleration Response with ±25% Variation 95

Figure 4.16 Interaction Plot for Lateral Acceleration Response with ±5% Variation 96

Figure 4.17 Full-Order SFEM Global Sensitivities with ±25% Parameter Variance 98

Figure 4.18 Convergence of Surrogate-based Global Sensitivity Indices for System Finite Element Model with ±25% variation on input parameters 100

Figure 4.19 Acceleration Time Histories for RVCH lateral motion 103

Figure 4.20 Reactor Vessel Closure Head with Simplified Head Assembly 105

Figure 4.21 Typical Pressurized Water Reactor Schematic 106

Figure 4.22 Acceleration Response Spectra Overlay of RVCH Lateral Motion 107

Figure 4.23 Contour Plot of Acceleration Response Spectra (acceleration magnitude represented by color scale) 107

Figure 4.24 Global Sensitivity Indices for System Finite Element Model Response Spectra... 109

Figure 4.25 Convergence of Surrogate-based Global Sensitivity Indices for System Finite Element Model Acceleration Response Spectra 111

Figure 4.26 Verification of Kriging Surrogate for SFEM with ±5% Input Parameter Variation 113

Figure 4.27 Verification of Kriging Surrogate for SFEM with ±25% Input Parameter Variation 114

Figure 4.28 Histogram of SFEM Output P17 with Varying Number of Training Samples 114

Figure 4.29 Overlay Comparison of Surrogate Verification RRMS Error and GSA Error for LHS 117

Figure 4.30 Overlay Comparison of Surrogate Verification RRMS Error and GSA Error for LPSS 117

Figure 5.1 RMS Error Metric Comparison between Verification and Cross-Validation Points 125

Figure 5.2 Histogram of Percent Differences between Cross-Validation and Verification Point Errors..... 126

Figure 5.3 Comparison of Latin Hypercube and Adaptive Refinement Samples for P4..... 129

Figure 5.4 Comparison of Latin Hypercube and Adaptive Refinement Samples for P10..... 129

Figure 5.5 Comparison of Latin Hypercube and Adaptive Refinement Samples for P4 and P10 130

Figure 5.6 Spearman Rank Correlation Coefficient Global Sensitivities for SFEM 134

Figure 5.7 Scatter Plot between Vertical Gap at Location 5 (P11) and RVCH Acceleration ... 135

Figure 5.8 Scatter Plot between Number of Components at Location 6 (P12) and RVCH Acceleration 135

Figure 5.9 Scatter Plot between Stiffness at Location 9 (P15) and RVCH Acceleration..... 136

Figure 5.10 Comparison of Global Sensitivities between Variance-based and Correlation Methods..... 136

Figure 6.1 Illustration of Typical Reactor Internals Assembly with Potential Applications for Inverse Mechanics Problems involving Sensitivity Analysis 141

Figure 6.2 Jennings' Method Oscillator Model (Nigam & Jennings, 1969) 147

Nomenclature

ASME	American Society of Mechanical Engineers
APDL	ANSYS[1] Parametric Design Language
DOE	Design of Experiment
FAST	Fourier Amplitude Sensitivity Test
FEA	Finite Element Analysis
FEM	Finite Element Method
FIV	Flow-Induced Vibration
GP	Gaussian Process
GSA	Global Sensitivity Analysis
LHS	Latin Hypercube Sampling
LOCA	Loss of Coolant Accident
LPSS	Latinized Partially Stratified Sampling
NSSS	Nuclear Steam Supply System
PWR	Pressurized Water Reactor
RPV	Reactor Pressure Vessel
RVCH	Reactor Vessel Closure Head
RVI	Reactor Vessel Internals
SFEM	System Finite Element Model
UI	Upper Internals

[1] ANSYS Mechanical, ANSYS Workbench and any and all ANSYS, Inc. brand, product, service, and feature names, logos, and slogans are trademarks or registered trademarks of ANSYS, Inc. or its subsidiaries located in the United States or other countries.

Preface

Computational modeling and simulation (M&S) provides a cost-effective means of answering difficult engineering questions. Testing is often cost prohibitive, and analytical solutions are often impractical to solve for complex problems, thus rendering computational modeling tools attractive. Moreover, computational models offer a means to diagnose (and prognosticate) the health of an engineering component or system. Thus, computational modeling plays a meaningful role in the overall task of engineering which ultimately contributes to the health and welfare of real people. Clean, reliable, and safe nuclear energy is one means of promoting such welfare.

Philosophers write about the "concretization of the ideal," in which otherwise idealistic notions are made concrete when put into practice. To some extent, this dissertation aims to make concrete things that many practitioners may consider to be merely idealistic notions. In fact, the methods employed within this dissertation are ripe for "concretization" in the context of engineering mechanics for nuclear power plants.

Computational models are imperfect solutions of sets of differential equations, which themselves are incomplete representations of reality. Furthermore, computational models carry with them certain pains. For example, computational fluid dynamics models are heavily dependent upon the turbulence model chosen to resolve the flow regime, requiring extensive experimental benchmarking to demonstrate validity. Random vibration of large systems likewise requires benchmarking of structural modes. Such dynamic problems take significant computational time to solve, and once one numerical solution is obtained, little is known regarding the uncertainties and meaningful parameters that influence input/output relationships.

I desire to make a contribution to ease this pain by concretely demonstrating the application of advanced methods (i.e., the "ideal") to the computational modeling and simulation of nuclear reactor stochastic dynamics. My aim is to do something for which the result permits engineering practitioners to:

1. Solve large dynamic problems in ways that are more computationally efficient than standard industrial practice.
2. Effectively gain understanding of uncertainties and meaningful parameters that may not otherwise be apparent.

The topics of Surrogate modeling and sensitivity analysis address both of these goals. That is, most dynamical systems can be characterized by some mathematical combination of their constituent parts (i.e., forcing functions, material properties, boundary conditions). If those constituent parts are known then they can be used to solve numerical models in ways that are less computationally burdensome and more physically insightful to the practicing engineer, thus permitting movement towards more efficient simulations for nuclear power plants.

The work communicated by this dissertation gratefully acknowledges the following individuals:

- to John Brigham of the University of Pittsburgh and Durham University, who patiently taught me computational mechanics and advised the associated work explored herein
- to Michael Shields of Johns Hopkins University, who patiently advised the use of stochastic methods explored herein
- to Gregory Meyer, Scott Sidener, and Richard Schwirian of Westinghouse who sacrificed significant time in teaching me the engineering practices of flow-induced vibration, probabilistic analysis, and acoustics, applied to the nuclear industry
- to Michael Tronosky and Matthew Kelley of Westinghouse, who advocated for my continued studies at the University of Pittsburgh
- to various colleagues at ETH-Zurich and ANSYS, who answered my many questions in using various software tools
- to the engineering faculty of the Ohio University who diligently coached me through the fundamentals of engineering towards the objective of creating something for good
- to my parents, Dan & Joan, my sister Laurie, and my brother* Adam, for shaping me into the man I am, who (for some reason) finds enjoyment in difficult things like listening to jazz music, running long distances, and writing a PhD dissertation
- to my wife, Jennifer, and my sons, Garrett, Grady, Joel, and Joshua, for patient endurance, for being a mosaic of beauty, joy, and chaos, and the motivation to excel in my vocation
- to the risen Jesus of Nazareth, who amidst our striving simply his disciples to breakfast

Thank you.

1.0 Introduction

Systems, structures, and components (SSCs) key to the safe and reliable operation of nuclear power plants (NPP) experience non-negligible random flow-induced vibration (FIV) loading during normal operating conditions. Furthermore, in accordance with the design basis for all operating plants, these same SSCs must be designed to withstand earthquakes and plant transients such as those associated with loss of coolant accidents (LOCA). These types of loading conditions may be considered stochastic processes. Quantification and propagation of uncertainties associated with characterizing the resultant structural dynamics, and therefore SSC reliability, constitutes a necessary but challenging endeavor.

The commercial nuclear industry is faced with significant cost challenges, in terms of both new plant designs and aging management of the operating fleet (Nuclear Energy Institute, 2014). The status quo is becoming increasingly unacceptable in that new plant designs experience cost-overruns and operating plants experience high maintenance costs to care for degradation associated with material reliability. Indeed, the future viability of nuclear energy hinges upon the ability of the industry to minimize cost while maintaining (or exceeding) current safety performance. Such costs have been noted to depend upon shifting project management practices to maximize the amount of design work performed prior to construction, as well as the promise offered by advanced reactor technology (i.e., passive safety features, small modular reactors, liquid metal reactors, etc.) (MIT Energy Initiative, 2018). Because engineering mechanics plays a central role in the design of a nuclear plant, methods that permit practitioners to enhance the credibility of modeling and simulation (M&S) of mechanics-based predictions

(Oberkampf, et al., 2007), while minimizing engineering costs (i.e., man-hours and/or computational resources) are of acute interest.

Uncertainty pervades engineering processes for nuclear plants, which span across many disciplines such as probabilistic risk assessment (PRA) and companion safety analysis, up-front component design for a new plant (i.e., forcing function development, sub-scale and start-up testing), the aging management of operating plants (i.e., stress corrosion cracking, fracture mechanics, and non-destructive examinations (NDE)), to the design of fuel (i.e., core loading pattern optimization, departure from nucleate boiling (DNB) correlations). Such uncertainties are sometimes easily correlated with one another, but often are not, and could be classified as aleatory in some instances and epistemic in others. Epistemic uncertainties could include that associated with having very small data sets from limited component inspections, while aleatory uncertainties could include the characterization of random turbulence in a particular flow domain. Furthermore, the contribution of uncertainties in design parameters to key outputs of interest is largely unknown in the early stages of design. Key outputs may include anything from stress intensities necessary to satisfy allowable limits per (ASME, 2017) to component reliabilities that directly influence core damage frequencies determined in PRA. This ill-posed portrait causes the design process to iterate, and thus take longer. For portions of new plants, construction often begins prior to design finalization, and therefore, in-field design changes occur which introduce very high construction costs. That is, an inefficient design process does not permit constructors the ability to await design finalization to begin construction for all systems of a nuclear plant. For an operating plant, lack of understanding of parameters can also manifest itself with increased maintenance costs. A better way forward, in which the contribution of parameter uncertainties may be quantified and adopted into design (and

companion aging management) processes is explored in this dissertation through the use of global sensitivity analysis (GSA) applied to stochastic dynamic applications associated with nuclear reactor structures. It is envisioned that if parameter uncertainties and the influence thereof can be properly understood early in the engineering analysis, then the (intertwined) design, analysis, construction, operations, and maintenance processes may become less iterative and thus more parsimonious.

In accordance with guidance provided by (United States Nuclear Regulatory Commission, 2017), new reactor designs are to complete a Comprehensive Vibration Assessment Program (CVAP) to evaluate FIV. License renewals similarly need to satisfy (United States Nuclear Regulatory Commission, 2017), but usually do so by demonstrating similarity to a valid prototype plant rather than completing a full analysis and test program, depending on the plant licensing basis. The CVAP includes extensive computational dynamic analysis, as described in (Westinghouse Electric Company, 2011), as well as a companion measurement and inspection program. The measurement and inspection program, as described in (Westinghouse Electric Company, 2015), recommends placing sensors in the locations that are dynamically correlated to the component response quantity of interest, which is generally near locations of high cyclic stress intensity. Correspondingly, the vibration analysis program establishes expected measurements (predictions) and associated acceptance criteria, which are based on the material fatigue life. Extensive verification and empirical validation of numerical models, for both forcing functions (e.g., turbulence and acoustic phenomena) and structural response, are employed to develop the predictions (Palamara, et al., 2015). The vibration analysis program thus requires the construction, verification, and validation of multiple

engineering simulations covering disciplines of both fluid and structural mechanics, and the interactions thereof.

One particular challenge to such analyses lies in the quantification and propagation of uncertainty of forcing functions associated with random turbulence. Flow within a nuclear reactor coolant system is highly turbulent. The high turbulence is necessary for core cooling and heat transfer, but also creates a substantial vibratory forcing function on the associated mechanical components and assemblies. For analysis of reactor internals, M.K. Au-Yang developed methods for determining forcing functions in a downcomer annulus (Au-Yang & Jordan, 1980), modeling random vibration induced by turbulent flow (Au-Yang & Connelly, 1977), and summarized the majority of his published work in (Au-Yang, 2001). Guidelines for practice are then based on the work of Au-Yang and others for dynamic analysis of nuclear components subjected to FIV (ASME, 2017). Recent industry efforts such as (Banyay, et al., 2015) have sought to improve the methods described in (ASME, 2017) and part of the aim of this work is directed toward supporting such industry efforts to promote the implementation of state-of-the-art analysis methods which can both lower computational costs and enhance the state of knowledge provided by dynamic analysis.

The forcing functions from turbulent flow acting on the various components of a nuclear reactor assembly can be characterized as a stochastic process (Tennekes & Lumley, 1972). As such, the dynamic structural response to turbulent loading is correspondingly random (Blevins, 2001), and accounting for this randomness is of paramount importance to proper system modeling. The spectral shape (i.e., the non-dimensional power spectral density, PSD) of the forcing random process may be known from prior test data of similar components in similar plants, but a best-estimate of the forcing function amplitude is generally difficult to determine.

Note that within the nuclear industry, the term “best estimate” analysis has been defined as an analysis that is “free of deliberate pessimism regarding selected acceptance criteria” and “includes uncertainty analysis” (International Atomic Energy Agency, 2008). Therefore, in practice, plant designers use biased approaches that bound the scatter in the available data to characterize forcing functions. In so doing, this may produce conservative design margins, but can result in over-designed equipment and contribute to a lack of understanding of the actual structural dynamic behavior of critical components. Furthermore, such approaches can lead to misleading conclusions regarding the true design margin of a given component or system; contributing to a false sense of confidence in a high margin (Type II error) or a false impression that a component has a low margin (i.e., is at risk of failure, Type I error). Driven largely by recent major structural failures caused by FIV in the nuclear industry, such as the steam generators at San Onofre (United States Nuclear Regulatory Commission, 2015) or the steam dryer at Quad Cities (United States Nuclear Regulatory Commission, 2013), recent revisions to the NRC regulatory guide (United States Nuclear Regulatory Commission, 2017) increasingly demand that bias and uncertainty be accounted for in prototype reactor designs. Note that the industry guidelines offer some flexibility in the sense that they do not necessarily require rigorous uncertainty quantification but rather require that uncertainty and bias have been accounted for in the design process. For example, regulatory guidance accepts an extensive uncertainty analysis such as described in (ASME, 2009), but regulators also accept a more simple uncertainty analysis, which demonstrates adequate bias to ensure conservative margins. Although the preceding discussion focuses primarily on FIV loads during normal operating conditions, quantification and propagation of uncertainties associated with modeling and simulation for safety analysis (i.e., LOCA) are likewise necessary (Kaizer, 2013).

Within a reactor assembly, multiple forcing functions are present that are attributed to different excitation mechanisms that have to do with component geometry, local coolant velocities and temperature, spectral shape functions, and overall plant configuration (i.e., reactor coolant pump or piping design). These individual fluid-borne forcing functions coupled with the structural dynamics, which depend on yet other uncertain model parameters such as joint stiffness, damping, and gap interface dimensions, constitute the total observed structural response. From this total response, it is often unclear which particular forcing functions or model parameters are the most relevant for the response of a given component. The effect of the different forcing functions and model parameters on the vibration response of a given component can vary significantly, with some having a relatively large effect and others being practically negligible. For example, the forcing function acting on an adjacent assembly (e.g., core barrel shell) may be more influential to the response of a given component (e.g., support columns or lower support structure) than the forcing function acting directly on that component. Misunderstanding the contributions of certain phenomenological behavior (e.g., the relative contribution of loads) can lead to excessive expense, such that analysts expend resources answering the wrong questions. For example, misunderstanding the contribution of boundary conditions and forcing functions to structural dynamic models may lead to studies employing advanced Computational Fluid Dynamics (CFD) models to accurately characterize forcing functions that do not meaningfully influence the structural response. Sensitivity analysis can, in part, serve to reduce this excessive effort and cost (Saltelli, et al., 2008). For example, sensitivity analysis in this context could inform the selection of the parameters to which model outputs are most sensitive. The ability to down-select parameters can inform where analytical or

experimental investment may be warranted to reduce epistemic uncertainties, or where real-time instrumentation and controls might focus to optimize plant operations and maintenance.

While local methods of sensitivity analysis provide only limited insight as to the influence of one variable while the others are fixed, and correlation-based GSA may only offer insight as to linearity or monotonicity of multiple terms, variance-based GSA provides a means of quantifying influence contributed by multiple uncertain parameters to changes (i.e., variance, scatter, "chaos") in key results of interest. Variance-based GSA, such as described in (Saltelli, et al., 2008), (Gratiet, et al., 2016), (Cannavo, 2012), and (Pianosi, et al., 2015), requires running a model a large number of times to properly characterize the relative importance of the various uncertain parameters (Schenk & Schueller, 2005). It is therefore of interest to explore options to either reduce the number of full-order runs (i.e., finite element analysis realizations) required to characterize sensitivity, or to altogether replace the full-order model with a surrogate model, which runs with minimal computational expense yet captures the relevant trends in the physical model. Surrogate modeling techniques aim to model physical phenomena using some means other than direct solution of the standard equations for the system physics (e.g., partial differential equations). This is useful when solving the governing equations is computationally expensive or when the governing equations are not known and the trends must be inferred from experimental data. Furthermore, several works have already applied the use of surrogate models for sensitivity analysis, successfully reducing the computational expense significantly such as (Shahsavani & Grimvall, 2011), (Hou, et al., 2016), (Cheng, et al., 2017).

Figure 1.1 shows an influence diagram used to construct an illustrative simplified cost-benefit analysis. In this postulated scenario, assumed cost quantities are shown in Table 1.1. In the influence diagram, the "Analysis Method" parallelogram represents an index associated with

three different approaches that may be taken to solve this problem. The "Number of Design Iterations" trapezoid is used to simulate how, as a design matures through time and thus iterates, the different analysis methods trend in terms of a cost comparison (represented by the hexagon). Design iterations are intended to represent a situation in which the physical geometry of a component changes, or a particular forcing function changes (perhaps due to an enhanced state of knowledge) and thus the dynamic response and corresponding ASME Code margins (i.e., design basis) change. In practice, both of these types of design iterations occur often; Chapter 4.0 considers the former situation (physical parameters), and Chapter 3.0 considers the latter situation (forcing functions). The study presented in Chapter 3.0 performs sensitivity studies on a fixed geometry and thus does not consider changes in geometric parameters, so the surrogates built and employed herein would not necessarily be valid in the context of geometry changes. On the other hand, the study presented in Chapter 4.0 performs sensitivity studies on how changes in model parameters such as geometry, stiffness, and mass can influence key results while holding the forcing function magnitude constant. Figure 1.2 then shows the results of how the engineering cost steadily increases when exercising a full-order model over many design iterations versus either supplementing a full-order model simulation with sensitivity analysis or by using a surrogate model in lieu of the full-order model (i.e., the items represented by the "Analysis Method" index). That is, significant savings may be realized by simply using the knowledge gained through sensitivity analysis to reduce complexity (i.e., the number of applied forcing functions) in the full-order model used for subsequent analyses in the design process; this is shown on the blue line in Figure 1.2. Further cost savings may be realized by altogether using a validated surrogate model in lieu of the full-order model in subsequent design iterations; this is shown on the green line in Figure 1.2. An important consideration is that the veracity of a

surrogate model requires that the parameter changes associated with the design iterations fall within the range of the parameter space sampled during initial model training. Therefore, there is a balance to achieve in which the sampled parameter space should be sufficiently large so as to bound the (future) parameter changes, but not so large so as to render the initial problem computationally prohibitive. In short, if some initial investment is made upfront to understand parameter sensitivity for large dynamic models, then significant cost savings can be realized as the design iterates and subsequent simulations are required.

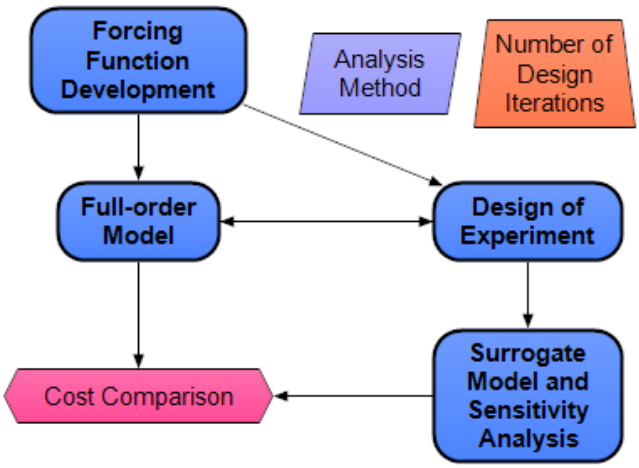

Figure 1.1 Influence Diagram and Inputs for Postulated Cost-Benefit Analysis

Table 1.1 Cost Benefit Analysis Variables

Variable	Value	Comment
Number of Design Iterations	10	Through the design and analysis of nuclear reactor structures, design parameters evolve, which warrants multiple iterations of computational models.
Cost of Full-order Model (per run)	10 hours	Finite element models subjected to dynamic FIV loading.
Cost of Initial Forcing Function Development (per function)	100 hours	Cost to run CFD and/or analyze test data to characterize forcing function. (e.g., boundary layer turbulence in downcomer annulus)
Number of Forcing Function Inputs	10	Multiple forcing functions acing on an assembly.
Number of runs needed for Design of Experiment	200	Number of full-order runs required for Surrogate model construction and/or GSA.
Cost of Surrogate Model construction	10 hours	Time required to develop and train Surrogate model.
Cost of Surrogate Model run	0.1 hours	Representation of the small cost of Surrogate model runs (it is recognized that actual Surrogate model cost is well below 0.1 hours).

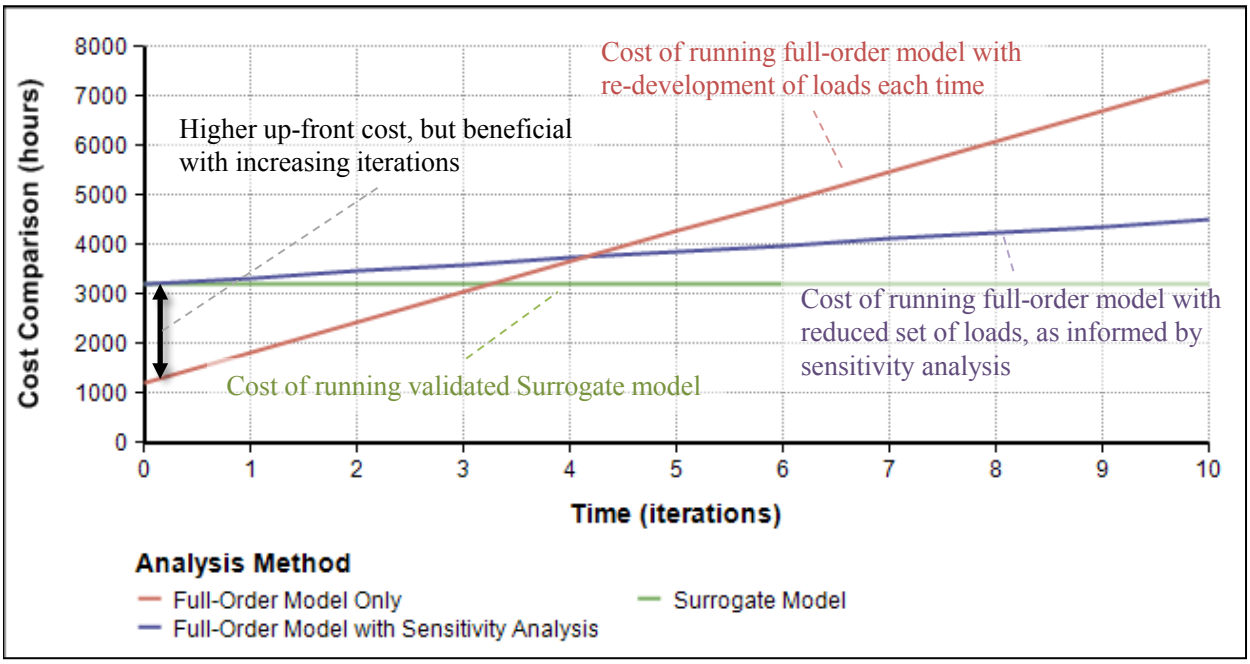

Figure 1.2 Illustration of Cost-Benefit Analysis between Traditional Approach and the use of Sensitivity Analysis

In this work, a framework is developed to use surrogate models for sensitivity analysis for random vibrations in a set of nuclear reactor internals, specifically to understand how forcing functions acting in different locations and mechanical model parameters influence the response of the structure. This is intended to support fast turnaround analyses for experimental (real-time or near real-time) plant diagnostics, operational prognostics, and component design in the presence of parameter uncertainties by informing decisions pertaining to the relative contributions of potentially contributing parameters. This research does not necessarily seek an optimal surrogate modeling technique, but rather a robust surrogate modeling technique that fits the specific application. For this purpose, Kriging (or Gaussian process regression) (e.g., (Nechak, et al., 2015) and (Huang, et al., 2011)) was used in this work. First, a full-order finite element model is used directly (i.e., without any complementary surrogate-model) to produce global sensitivity indices which do not meaningfully change upon generation of further full-order model realizations. By first computing sensitivity indices from full-order runs, this provides a set of results against which the surrogate-computed sensitivity indices may be compared. Given these "converged" sensitivity indices, the sampling of the parameter space for the full-order model runs is then investigated by considering a computational design of experiment (DOE) using Latin Hypercube (LHS) and Latinized Partially Stratified sampling (LPSS) techniques. Thus stability is evaluated in terms of the GSA result when the Kriging surrogate is trained by different sampling as a means of providing confidence in the surrogate-based sensitivities. One aim of this is to identify if it is practical to minimize the number of full-order runs required to re-generate sensitivity indices by way of a trained Surrogate model. An engineering design organization could thus expect that significant economic benefit may be realized in the engineering design process of nuclear reactor structures by coupling the advantages offered by

both Surrogate modeling and GSA. The other aim, in this research, is to examine the relationship of the sensitivity indices to metrics in surrogate model validation. This is novel in the sense that observation of the error of the sensitivity indices relative to a baseline provides insight as to the veracity of the Surrogate model with which the sensitivity indices are computed.

1.1 Dissertation Outline and Research Questions Addressed

A brief outline of this dissertation is provided in Table 1.2.

Table 1.2 Outline for Dissertation

Chapter		Topics Addressed
1.0	Introduction	Motivation for research by reviewing topics pertinent to nuclear reactor structural dynamics, and providing overview of Surrogate modeling and sensitivity analysis.
2.0	Methodology	Details provided for approaches to Random Vibration, Global Sensitivity Analysis, Surrogate Modeling
3.0	Application to Stationary Random Vibration for Upper Internals Assembly	Finite element model described, Full-order model based Sensitivity analysis, Surrogated-based sensitivity analysis based on Fourier Amplitude Sensitivity Test (FAST), LHS, and LPSS Design of Experiments (DOE), quantification of Surrogate verification error and Surrogate-based GSA accuracy.
4.0	Application to Non-Stationary Random Vibration of Reactor System	Finite element model described, Full-order model based Sensitivity analysis, Surrogated-based sensitivity analysis based on LHS, and LPSS DOEs, quantification of Surrogate verification error and Surrogate-based GSA accuracy. Additional examinations to the effects imparted by increasing parameter variation, and comparing effect of spectral versus transient results.
5.0	Effectiveness of Surrogate Modeling Methodological Changes	Using the system finite element model (SFEM) from Chapter 4.0 , introduce the effectiveness of methodological changes associated with an aggregate surrogate modeling technique and adaptive sampling.
6.0	Conclusions of Dissertation	Summary of research contributions and recommendations for building upon this work.

Table 1.3 summarizes the questions sought to explore by this research, and provides an outline of where responses to these questions may be found in this dissertation.

Table 1.3 Research Questions

Question	**Response**
Can a Kriging surrogate accurately represent an Upper Internals (UI) model subjected to stationary FIV and a SFEM subjected to non-stationary loading?	Yes, see Sections 3.3.1 and 4.2.4
Does the sampling method impart a significant effect on the accuracy of the surrogate model, or the global sensitivities computed from the surrogate model for either application?	Yes at small sample sizes, but not at larger sample sizes, see Sections 3.3.2 and 4.2.2
How does the surrogate model accuracy change, and does the effectiveness of LPSS become more pronounced, when the input parameter variance is increased?	See Section 4.2.3.1
How do the sensitivities associated with acceleration response spectra change when evaluating non-stationary loading of the SFEM?	See Section 4.2.3.2
How does the use of an aggregate surrogate model improve the verification error?	See Section 5.1
How does the use of adaptive sampling improve the verification error?	See Section 5.2
How might correlation versus variance-based methods of GSA differ?	See Section 5.3

2.0 Methodology

In the studies presented herein, finite element analysis serves as the "full-order model" of the system of interest, which is used to generate the baseline sensitivity indices, as well as to create the datasets to train the surrogate models – from which sensitivity indices will be likewise computed and verified against those from the full-order model. The overall methodology used seeks to combine the construction of a computational DOE, sensitivity analysis, and surrogate modeling as illustrated by the three parallel workflows in Figure 2.1. In the first workflow, GSA is performed on the full-order model using the Fourier Amplitude Sensitivity Test (FAST) algorithm as one means of computing sensitivity indices. This is done to establish the "true" sensitivity indices to compare with those computed by exercising a surrogate model. Using these full-order model evaluations from the FAST GSA, surrogate models for different numbers of realizations from the finite element analysis are constructed (for the UI FIV model). In workflows 2 and 3, the full-order model is run using two different methods for computational DOE as an alternate means of building a surrogate model for GSA. The three workflows are then compared to evaluate the effectiveness of these different sampling methods for use in the dynamic analysis of a nuclear reactor assembly as well as the stability of the GSA results computed from the various surrogates.

The models evaluated within this document constitute simulations of random vibration, for which the methodology is discussed in Section 2.1. One desires to compute sensitivities from those models, for which the methodology is discussed in Section 2.2. To calculate those sensitivities in a computationally efficient manner, surrogate models are constructed, for which the methodology is discussed in Section 2.3. Finally, to construct surrogate models, a training

data set is necessary and so sampling strategies for developing such a training set are discussed in Section 2.4.

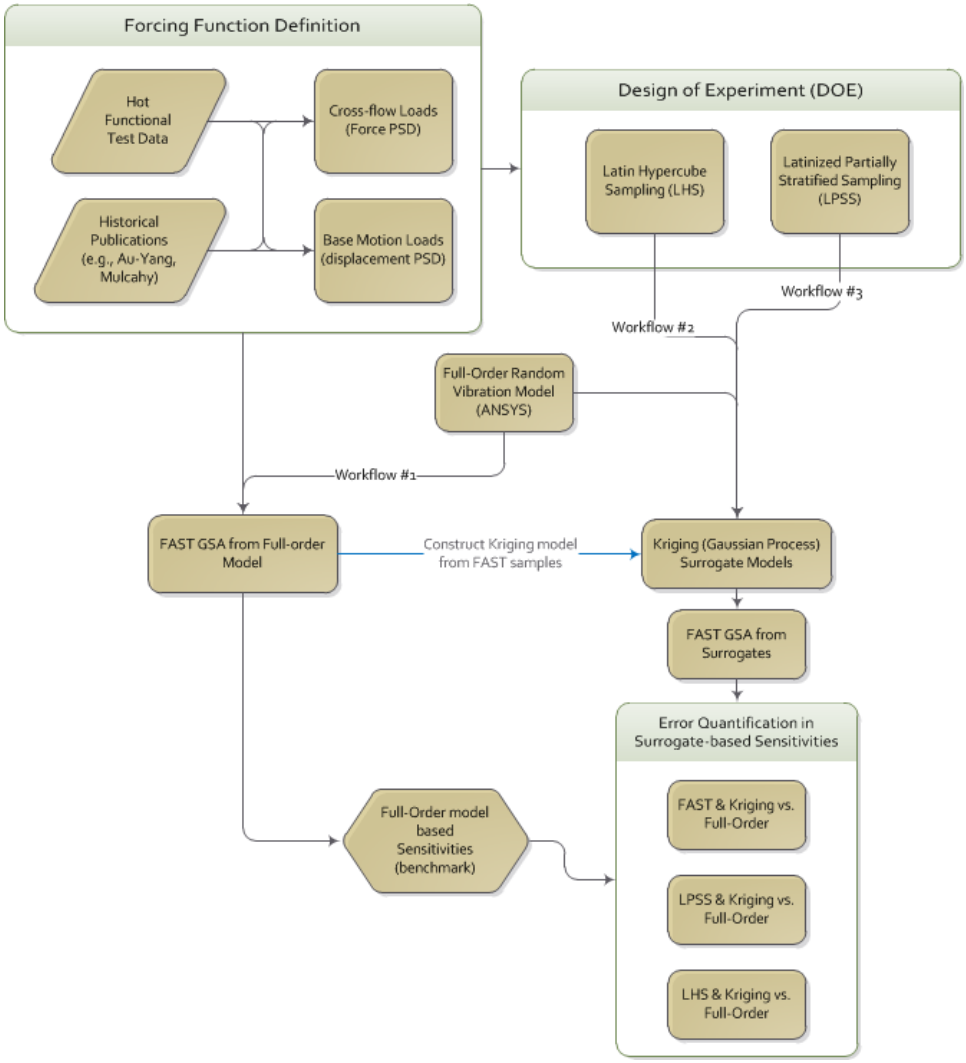

Figure 2.1 Flow Chart of Analytical Workflow involving model definition, Computational DOE, Surrogate Training, and GSA

2.1 Random Vibration

2.1.1 Stationary Random Vibration

Flow-induced random vibrations in reactor internals can often be characterized as stationary, ergodic random processes ((Tennekes & Lumley, 1972), (Blevins, 2001)). The stationary property follows from the steady-state operating conditions under which nuclear reactors typically operate; hence the vibration characteristics do not change with time. Thanks to the ergodic property, the mean and autocorrelation function of the forcing process, $z(t)$, can be defined, respectively, from a single realization as (Bendat & Piersol, 2010):

$$\mu = \lim_{T\to\infty} \frac{1}{T} \int_0^T z(t)\, dt \tag{1}$$

$$R(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_0^T z(t) z(t+\tau)\, dt \tag{2}$$

where τ is the time lag. The autocorrelation function may be thought of as a correlation between the values of the random process at two different times, t ands $t+\tau$. Physically, τ should be chosen as a sufficiently small time lag as to resolve the highest frequencies of interest, which may pertain to the time scale of the dominant turbulent eddies or the dynamics of the structure itself.

The forcing random process (e.g., applied force and displacement loading) can be transformed to the frequency domain, with the power spectral density (PSD) function given by the Wiener-Khintchine transform of the autocorrelation function:

$$G(\omega) = \int_{-\infty}^{\infty} R(\tau) e^{-\omega t}\, d\tau \tag{3}$$

This G PSD term is used to define the forcing function in the finite element problem, which is correspondingly solved in the frequency domain across a frequency range adequate to encompass at least 95% of the dominant vibration modes (e.g., 0 to 500 Hz). The finite element method with modal superposition was used herein to solve for the system output PSDs (specifically, the commercial finite element analysis software ANSYS was used (ANSYS, 2016)). In particular, the natural frequencies, ω, and corresponding mode shapes, ϕ, of the structure were computed using the Lanczos algorithm (Rajakumar & Rogers, 1991). Then, the single degree-of-freedom (SDOF) transfer functions were used to calculate the response PSDs, as briefly outlined below.

The full equation of motion solved herein by the finite element method may be expressed in terms of displacement u as a function of free (f) and restrained (r) mass, M, stiffness, K, and damping, C, matrices subjected to forcing function, F:

$$\begin{bmatrix} [M_{ff}] & [M_{fr}] \\ [M_{rf}] & [M_{rr}] \end{bmatrix} \begin{Bmatrix} \ddot{u}_f \\ \ddot{u}_r \end{Bmatrix} + \begin{bmatrix} [C_{ff}] & [C_{fr}] \\ [C_{rf}] & [C_{rr}] \end{bmatrix} \begin{Bmatrix} \dot{u}_f \\ \dot{u}_r \end{Bmatrix} + \begin{bmatrix} [K_{ff}] & [K_{fr}] \\ [K_{rf}] & [K_{rr}] \end{bmatrix} \begin{Bmatrix} u_f \\ u_r \end{Bmatrix} = \begin{Bmatrix} \{F\} \\ \{0\} \end{Bmatrix} \tag{4}$$

Projecting the governing equations onto the system mode shapes, the equation of motion for modal dynamics of which a solution is sought may be expressed as:

$$\ddot{y}_j + 2\zeta_j \omega_j \dot{y}_j + \omega_j^2 y_j = \Lambda_j \tag{5}$$

where j is the mode number (from 1 to n), and y, and ζ, are the corresonding generalized displacement amplitude, and damping ratio for each of the modes, and Λ is the modal load. Note that ω_j is a system property (i.e., the square root of stiffness of the j^{th} mode divided by mass of the j^{th} mode). For the j^{th} mode, the modal loads are defined in terms of modal participation factors for nodal excitation γ_j as:

$$\Lambda_j = \gamma_j = \{\phi_j\}^T \{F\} \tag{6}$$

Ultimately, the FIV analysis done in Chapter 3.0 computes the mean squared axial strain response of the upper support skirt. Therefore, it was necessary to obtain the RMS response from a PSD analysis. Given the damping of the j^{th} mode, ζ_j, the corresponding SDOF transfer function, H_j, for an input force may be computed as:

$$H_j(\omega) = \frac{1}{\omega_j^2 - \omega^2 + i\left(2\zeta_j \omega_j \omega\right)} \tag{7}$$

For nodal force excitations, the dynamic and pseudo-static modal PSDs can then be expressed in terms of these transfer functions and mode superposition as:

$$\begin{aligned} \rho_{jk}(\omega) &= \sum_{l=1}^{r_1} \sum_{m=1}^{r_1} \gamma_{lj} \gamma_{mk} H_j^*(\omega) H_k(\omega) \bar{G}_{lm}(\omega) \\ \bar{\rho}_{lm}(\omega) &= \frac{1}{\omega^4} \hat{G}_{lm}(\omega) \end{aligned} \tag{8}$$

where γ_{lj} and γ_{mk} are the participation factors from the modal analysis for modes j and k, respectively, corresponding to force excitation l and m, respectively, and H^* indicates the complex conjugate of H. $\bar{G}_{lm}(\omega)$ and $\hat{G}_{lm}(\omega)$ represent the input force and acceleration PSDs, respectively, which are supplied to the finite element analysis through forcing function development in this case (i.e., pre-defined). The number of nodal and base PSDs is denoted by r_1 and r_2, respectively. Furthermore, for uncorrelated PSD inputs, $l = m$.

The PSD responses may thus be expressed in terms of the modal PSDs, ρ_{jk} and $\bar{\rho}_{lm}$, as dynamic and pseudo-static parts, respectively in terms of the modal strains and static strains as:

$$G_{d_i}(\omega) = \sum_{j=1}^{n} \sum_{k=1}^{n} \bar{\phi}_{ij} \bar{\phi}_{ik} \rho_{jk}(\omega) \tag{9}$$

$$G_{s_i}(\omega) = \sum_{l=1}^{r_2} \sum_{m=1}^{r_2} \bar{A}_{il} \bar{A}_{im} \bar{\rho}_{lm}(\omega)$$

Where modal and static strain terms $\bar{\phi}$ and $\bar{A}$, respectively, would be calculated using standard mesh-based differentiation and approximation from the associated displacement terms, ϕ and A. Finally, the mean square strain response (ϵ^2) may be expressed as:

$$\epsilon_{f_i}^2 = \int_0^\infty G_{d_i}(\omega) d\omega + \int_0^\infty G_{s_i}(\omega) d\omega \tag{10}$$

in which $\int_0^\infty G_{d_i}(\omega) d\omega$ is the variance of the i^{th} relative (dynamic) free strains, and $\int_0^\infty G_{s_i}(\omega) d\omega$ is the variance of the i^{th} pseudo-static strains; note that no covariance terms are included in the formulation used herein, and each forcing function was defined without correlation. A summary of the relevant theory of stochastic dynamics may be found in (Ortiz, et al., 1995).

2.1.2 Non-Stationary Random Vibration

Various transient scenarios for which a nuclear plant must be designed are non-stationary random processes, such as a loss of coolant accident (LOCA) or an earthquake. These events are non-stationary in that the mean value and frequency structure can be time varying. These events are typically simulated as time-histories through transient finite element analysis to estimate structural dynamic behavior for purposes of reactor component qualification.

2.1.2.1 Finite Element Formulation

The detailed finite element method is not presented here, as details may be found in (Reddy, 2006) or Section 6.2.4 of (ANSYS, 2016), but the basic governing equations are shown as follows. Starting from the discretized equation of motion as Equation (4), and adding time discretization by using subscript n to indicate time step dependence:

$$[M]\{\ddot{u}_{n+1}\} + [C]\{\dot{u}_{n+1}\} + [K]\{u_{n+1}\} = \{F_{n+1}\} \tag{11}$$

where $\ddot{u}_{n+1}$, $\dot{u}_{n+1}$, and u_{n+1} are the nodal acceleration, velocity, and displacement vectors at time t_{n+1}, and F_{n+1} is the associated applied load. Let $t_{n+1} = t_n + \Delta t$. The Newmark time integration scheme used herein, in which α and δ are the integration parameters per (Hughes, 1987), marches through time by defining the velocity and displacement vectors as:

$$\{\dot{u}_{n+1}\} = \{\dot{u}_n\} + [(1-\delta)\{\ddot{u}_n\} + \delta\{\ddot{u}_{n+1}\}]\Delta t \tag{12}$$

$$\{u_{n+1}\} = \{u_n\} + \{\dot{u}_n\}\Delta t + \left[\left(\frac{1}{2} - \alpha\right)\{\ddot{u}_n\} + \alpha\{\ddot{u}_{n+1}\}\right]\Delta t^2 \tag{13}$$

by combination of equations (11) through (13), a single-step time integration equation in terms of unknown $\{u_{n+1}\}$ and the known mass, stiffness, and damping matrices may be written as:

$$(a_0[M] + a_1[C] + [K])\{u_{n+1}\}$$
$$= \{F_{n+1}\} + [M](a_0\{u_n\} + a_2\{\dot{u}_n\} + a_3\{\ddot{u}_n\}) \quad (14)$$
$$+ [C](a_1\{u_n\} + a_4\{\dot{u}_n\} + a_5\{\ddot{u}_n\})$$

where the time integration terms are summarized in Table 2.1.

Table 2.1 Time Integration Coefficients used for Transient Finite Element Method

a_0	a_1	a_2	a_3	a_4	a_5
$\frac{1}{\alpha\,\Delta t^2}$	$\frac{\delta}{\alpha\,\Delta t}$	$\frac{1}{\alpha\,\Delta t}$	$\frac{1}{2\alpha} - 1$	$\frac{\delta}{\alpha} - 1$	$\frac{\Delta t}{2}\left(\frac{\delta}{\alpha} - 2\right)$

Thus, the displacement at the $n+1$ time step may be computed. The corresponding velocity and acceleration can then be calculated using:

$$\{\dot{u}_{n+1}\} = a_1(\{u_{n+1}\} - \{u_n\}) - a_4\{\dot{u}_n\} - a_5\{\ddot{u}_n\} \quad (15)$$

$$\{\ddot{u}_{n+1}\} = a_0(\{u_{n+1}\} - \{u_n\}) - a_2\{\dot{u}_n\} - a_3\{\ddot{u}_n\} \quad (16)$$

2.1.2.2 Acceleration Response Spectra

Given the transient solution of a finite element analysis subjected to nonstationary loading, it is often insightful to express the response as a function of frequency. In the case of non-stationary data, the autocorrelation function becomes not only a function of time lag, τ, but time, t, itself, since the mean value changes with time. As such, for non-stationary data, Equation (2) becomes:

$$R(t, t-\tau) = \lim_{T\to\infty} \frac{1}{T} \int_0^T z(t) z(t+\tau)\, dt \tag{17}$$

While the previously discussed PSD method works well for the simulation of stationary random vibration, the response spectrum method is primarily used for characterizing the frequency content of non-stationary random vibration within the nuclear industry. Approaches of formulating response spectra are vast, and a brief survey of noteworthy contributions to this topic is provided in Appendix B. A concise definition of a response spectrum is provided in Paragraph N-1110(d) of (ASME, 2017) as follows:

> "The response spectrum is defined as a plot of the maximum response (acceleration, velocity, or displacement) of a family of idealized linear single-degree-of-freedom damped oscillators as a function of natural frequencies (or periods) of the oscillators to a specified vibratory motion input at their supports."

In practice, response spectra are used in dynamic analyses of nuclear power plants per (United States Nuclear Regulatory Commission, 2014), and can be computed using such software as (Converse & Brady, 1992) (a Matlab implementation of this code is provided in Appendix B). In fact, the methodology for response spectra calculation laid out in (Westinghouse Electric Company, 1976) cites (Tsai, 1972), and both of those documents are cited in Appendix N of (ASME, 2017), in terms of pertinent codes and standards. To evaluate the response spectra, it is first necessary to consider the equation of motion at some arbitrary location for a simple oscillator:

$$\ddot{u} + \frac{c}{m}\dot{u} + \frac{k}{m}u = -\ddot{z} \tag{18}$$

If $\omega = \sqrt{\frac{k}{m}}$ is the undamped natural frequency and $\zeta = \frac{c}{2m\omega}$ is the percent of critical damping, then Equation (18) may be expressed as:

$$\ddot{y} + 2\zeta\omega\dot{y} + \omega^2 y = -\ddot{z} \tag{19}$$

Following the derivation of response spectra from this general equation of motion per (Nigam & Jennings, 1969), let the forcing acceleration be approximated as a segmentally linear function $-\ddot{z}_i - \frac{\Delta \ddot{z}_i}{\Delta t_i}(t - t_i)$, for which the solution to Equation (19) is given by:

$$y = e^{-\zeta\omega(t-t_i)}\left[C_1 \sin\omega\sqrt{1-\zeta^2}(t-t_i) + C_2 \cos\omega\sqrt{1-\zeta^2}(t-t_i)\right] - \frac{a_i}{\omega^2} + \frac{2\zeta}{\omega^3}\frac{\Delta \ddot{z}_i}{\Delta \mathrm{t_i}} - \frac{1}{\omega^2}\frac{\Delta \ddot{z}_i}{\Delta \mathrm{t_i}}(t - t_i) \tag{20}$$

where C_1 and C_2 are constants of integration given as:

$$C_1 = \frac{1}{\omega\sqrt{1-\zeta^2}}\left(\zeta\omega y_i + \dot{y}_i - \frac{2\zeta^2 - 1}{\omega^2}\frac{\Delta \ddot{z}_i}{\Delta t_i} + \frac{\zeta}{\omega}\ddot{z}_i\right) \tag{21}$$

$$C_2 = y_i - \frac{2\zeta}{\omega^3}\frac{\Delta \ddot{z}_i}{\Delta t_i} + \frac{\ddot{z}_i}{\omega^2} \tag{22}$$

Substituting C_1 and C_2 into Equation(20) yields:

$$\boldsymbol{y}_{i+1} = \boldsymbol{A}(\zeta, \omega, \Delta t_i)\boldsymbol{u}_i + \boldsymbol{B}(\zeta, \omega, \Delta t_i)\ddot{\boldsymbol{z}}_i \tag{23}$$

where $\boldsymbol{y_i} = \begin{Bmatrix} y_i \\ \dot{y}_i \end{Bmatrix}$, and variables $\boldsymbol{A}$ and $\boldsymbol{B}$ indicate matrix quantities which are detailed in (Nigam & Jennings, 1969).

As an alternative method of determining the response spectra, according to (Westinghouse Electric Company, 1976) a Laplace transform of Equation (19) provides an expression for the relative displacement (with respect to foundation) as a function of time:

$$u(t) = u(0)e^{-\alpha\omega t}\left(\cos(pt) + \frac{\zeta\sin(pt)}{\varrho}\right) + \dot{u}(0)\frac{e^{-\zeta\omega t}\sin(pt)}{p} - \frac{\int_0^t \ddot{z}(\gamma)e^{-\zeta\omega(t-\gamma)}\sin p(t-\gamma)\,d\gamma}{p} \tag{24}$$

where $p = \omega\varrho$ and $\varrho = \sqrt{1-\zeta^2}$. Given the time-dependent relative displacement from Equation(24), the absolute displacement and acceleration may be derived, and thus, a set of equations for the relative displacement and velocity at the n^{th} time interval can be determined.

Finally, whether using the method per (Nigam & Jennings, 1969) or (Westinghouse Electric Company, 1976), the absolute response acceleration at the i^{th} time interval is:

$$\ddot{u}_i + \ddot{z}_i = -(\omega^2 u_i + 2\zeta\omega\dot{u}_i) \tag{25}$$

Hence, if the displacement and velocity of the oscillator are known at some time t_0, the state of the oscillator at all subsequent times t_i can be computed by application of Equations (23) and (25). This calculation is repeated for each natural frequency of the dynamic system, and the absolute maximum acceleration at each frequency is plotted as a function of that frequency to produce the acceleration response spectrum.

2.2 Global Sensitivity Analysis

While many methods of sensitivity analysis exist (Morgan, et al., 1992), GSA is employed herein as a variance-based technique, which surveys the full parameter space by evaluating all values each parameter could have with respect to one another, considering the probability distribution defined for each parameter. As such, GSA appropriately handles situations in which multiple parameters could have different magnitudes, and the combined effect of those different magnitudes has some effect on the desired outcome; this may be contrasted with respect to varying parameters on a one-at-a-time basis. GSA provides insight as to the relative importance of multiple parameters (e.g., forcing functions or physical characteristics), which mutually influence the system response of interest. Furthermore, GSA can also account for uncertainty in the input parameter space, so that each plausible combination of relative parameter variations is considered.

First-order global sensitivity indices for output $y = f(\boldsymbol{p})$ given input parameters $\boldsymbol{p} = (p_1, p_2, \dots, p_n)$ are defined as (Saltelli, et al., 2008):

$$S_i = \frac{V[E(y|p_i)]}{V(y)} \tag{26}$$

where V[·] denotes the variance operator. The expected value of y can be evaluated by the n dimensional integral:

$$E(y) = \int_{I^n} f(\boldsymbol{p}) d\boldsymbol{p} \tag{27}$$

in which I^n is the n dimensional unit hypercube. The Fourier amplitude (FAST) method is used for the present study to approximate Equation (27) by converting the n dimensional integral into a one-dimensional integral as a function of a new variable s.

2.2.1 Fourier Amplitude Sensitivity Test (FAST) Method

The essence of FAST is to generate a curve in the parameter space that is a periodic function of each parameter, with a different frequency for each. The contribution of each input is measured by the contribution of its characteristic frequency Ω_i to the outputs (Morgan, et al., 1992). First, per (Cannavo, 2012) the function $f(\boldsymbol{p})$ may be expanded as:

$$f(\boldsymbol{p}) = f_0 + \sum_{s=1}^{n} \sum_{i_1<\cdots<i_s} f_{i_1\ldots i_s}(p_{i_1}, \ldots, p_{i_s}) \tag{28}$$

While the first summation is over the n input variables, the second summation is over all possible combinations of s different input variables, hence at most 2^n components. The function $f(\boldsymbol{p}) = f_{i_1\ldots i_s}(p_{i_1}, \ldots, p_{i_s})$ can then be expressed as a Fourier series:

$$f(\boldsymbol{p}) = \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \cdots \sum_{k_n=-\infty}^{\infty} C_{k_1 k_2 \ldots k_n} e^{j2\pi(k_1 p_1 + k_2 p_2 + \cdots + k_n p_n)} \tag{29}$$

With Fourier coefficients:

$$C_{k_1 k_2 \ldots k_n} = \int_{I^n} f(p) e^{j2\pi(k_1 p_1 + k_2 p_2 + \cdots + k_n p_n)} dp \tag{30}$$

The variances V of the function $f(p)$ equate to the sums of the parts of the Fourier coefficients:

$$V\{f_{i_1\ldots i_s}\} = \sum_{k_{i_1}=-\infty}^{\infty} \cdots \sum_{k_{i_s}=-\infty}^{\infty} \left| C_{k_{i_1}\ldots k_{i_s}} \right| \tag{31}$$

A multi-dimensional integral must be solved to calculate the Fourier coefficients of Equation (30). In order to solve the n dimensional integral, every input can be expressed as a function of a new independent variable s as:

$$p_i(s) = \frac{1}{2} + \sin^{-1}(\sin(\Omega_i s)) \tag{32}$$

where the set $\{\Omega_1, \dots, \Omega_n\}$ is linearly independent of integer frequencies. Applying the change of variable, the expected value of y can then be expressed as $\frac{1}{2\pi}\int_{-\pi}^{\pi} f(s)ds$. Solution of this integral for the expected value of y involves an analysis of variance (ANOVA) decomposition, which includes the calculation of the Fourier coefficients C_{k_i}. The Fourier coefficients can be estimated by numerical integration of:

$$C_{k_i} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(s)e^{-j2\pi k_i \Omega_i s}ds \tag{33}$$

One can compute the Fourier coefficients to some desired accuracy by selecting enough values of s, where M is some user-defined integer:

$$s = {}^{2\pi q}/_{M}, \quad q = 1,2,\dots,M \tag{34}$$

The number of discrete intervals used to evaluate this integral is thus defined by variable M. Per (McRae, et al., 1980), the choice of M and the number of inputs k govern the number of model runs used to compute the GSA indices. Therefore, as the number of inputs typically given for a problem of interest, the parameter M can be adjusted to control the number of model runs required to estimate the global sensitivity indices. Finally, the numerator of Equation (26) needed for computing the global sensitivity indices is calculated as:

$$E(y|p_i) = \sum_i |C_{k_i}| \tag{35}$$

Substituting Equation (35) into Equation (26) provides the first-order global sensitivity indices as:

$$S_i = \frac{V\left[\sum_i |C_{k_i}|\right]}{V(y)} \tag{36}$$

2.2.2 Correlation based Methods

Correlation based methods constitute one class of approaches for GSA which seek to provide insight as to the statistical dependence between variables. For example, the Spearman rank correlation method assesses how well the relationship between two variables can be explained using a monotonic function. In Chapter 5.0 , the GSA performed on SFEM briefly evaluates a Spearman rank correlation method in addition to FAST.

Given a set of uncertain inputs $x_1 x_2 \dots x_n$, and the corresponding outputs $y_1 y_2 \dots y_n$, the absolute value of the rank correlation provides a measure of the relative degree to which change in variable x_i contributes to change in output y_i. More negative rank correlation values indicate a strong negative monotonic tendency (i.e., output decreases while input increases) while more positive rank correlation values indicate a strong positive monotonic tendency (i.e., output increases while input increases). Two statistically independent quantities would have an expected rank correlation of zero.

Define the rank as the relative position label of the observations within the variable (e.g., 1st, 2nd, 3rd, etc.), so for the inputs as:

$$R^{in} = \left\{ r_i^{(j)} \in \{1, \dots, n\} : r_i^{(j)} > r_i^{(k)} \Leftrightarrow x_i^{(j)} > x_i^{(k)}, j \in \{1, \dots, n\} \right\} \tag{37}$$

The rank transformed model response then is $R^{out} = \left\{ r_y^{(1)}, r_y^{(2)}, \dots, r_y^{(n)} \right\}$. The Spearman rank-order correlation coefficient may be expressed as:

$$r_s = \frac{\sum_i^n \left(R_i^{in} - \bar{R}^{in}\right)\left(R_i^{out} - \bar{R}^{out}\right)}{\sqrt{\sum_i^n \left(R_i^{in} - \bar{R}^{in}\right)^2}\sqrt{\sum_i^n \left(R_i^{out} - \bar{R}^{out}\right)^2}} \tag{38}$$

Where:

R_i^{in} rank of x_i within the set of observations $[x_1 x_2 \dots x_n]^T$

R_i^{out} rank of y_i within the set of observations $[y_1 y_2 \dots y_n]^T$

$\bar{R}^{in}, \bar{R}^{out}$ average ranks, across all n observations, of R_i^{in} and R_i^{out} respectively

2.3 Surrogate Modeling

2.3.1 Gaussian Process Regression

In recent years, reduced-order modeling techniques (e.g. (Grigoriu, 2010), (Grigoriu & Field, 2014)) and surrogate modeling methods (e.g., (Paez, et al., 1997)) have gained popularity for random vibration problems. One popular surrogate modeling method is kriging, otherwise known as Gaussian process modeling or Gaussian process regression, which has been shown to be effective for stochastic structural dynamics. For example, (Abbiati, et al., 2017) successfully used Kriging in conjunction with hybrid simulation to establish an active learning method in the context of structural reliability analysis for seismic applications. Kriging has the advantage when compared to some other popular surrogate modeling strategies of providing an error metric in the variance of the surrogate model, and has been successfully studied alongside methods of sensitivity analysis (Gratiet, et al., 2016).

A kriging model, $\mathcal{M}^K$, serving as a surrogate for the full-order model $\mathcal{M}$ (e.g., the random vibration finite element model), is expressed in accordance with (Rasmussen & Williams, 2005), (Sacks, et al., 1989), and (Lataniotis, et al., 2017), in which the bold-faced variables indicate a vector quantity:

$$\mathcal{M}^K(\boldsymbol{x}) = \boldsymbol{\Psi}(\boldsymbol{x}) + \sigma^2 \boldsymbol{R}(\boldsymbol{x}; \theta) \tag{39}$$

in which $\boldsymbol{\Psi}(\boldsymbol{x}) = \boldsymbol{f^T}(\boldsymbol{x})\boldsymbol{\beta}$ is the mean value (or trend) constructed from regression coefficients $\boldsymbol{\beta}$ and basis functions $\boldsymbol{f}(\boldsymbol{x})$, which is computed as the summation:

$$\boldsymbol{\Psi}(\boldsymbol{x}) = \sum_{\alpha \in \boldsymbol{A}^{M,P}} \beta_\alpha f_\alpha(\boldsymbol{x}) \tag{40}$$

where the basis functions were taken as multivariate polynomials of the form $f_{\alpha}(\boldsymbol{x}) = \prod_{i=1}^{M} x^{\alpha_i}$, $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \dots, \alpha_M\}$ is a vector of indices, and $\boldsymbol{A}^{M,P} = \{\boldsymbol{\alpha}: |\boldsymbol{\alpha}| \leq P\}$ that yield polynomials in the M input variables up to degree P. For this work, ordinary Kriging was used in which the mean (trend) had a constant yet unknown value, which may be simply expressed as $\boldsymbol{f^T}(\boldsymbol{x})\beta = f_1(x)\beta_1 = \beta_1$ (Marelli & Sudret, 2014), in which case $P = 1$.

The second term in Eq. (39) $\sigma^2 R(\boldsymbol{x}; \theta)$ is a zero-mean stationary Gaussian random process with variance σ^2 and autocorrelation function $R(\boldsymbol{x}; \boldsymbol{\theta})$ with which a correlation matrix $\boldsymbol{R}$ is populated. For this work, $R(\boldsymbol{x}; \boldsymbol{\theta}) = R(x, x'; \boldsymbol{\theta})$ is an n-dimensional separable ellipsoidal correlation function expressed as:

$$R(x, x'; \theta) = R(h), \qquad \text{where} \qquad \boldsymbol{h} = \sqrt{\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \left(\frac{\boldsymbol{x_i} - \boldsymbol{x_i}'}{\boldsymbol{\theta_i}} \right)^{\boldsymbol{2}}} \tag{41}$$

where the parameter θ is determined using maximum likelihood estimation (MLE).

MLE is used with a specified trend and correlation function, and a Kriging surrogate is established by estimating parameters β (for trend), and σ^2, θ (for correlation) which maximize the likelihood of realizing the actual (known) function evaluations (i.e., from full-order model $\mathcal{M}$). The operations by which MLE is used to perform this calculation are detailed in (Lataniotis, et al., 2017) or (Sundar & Shields, 2018).

2.3.2 Genetic Aggregation

Although a surrogate model such as the Kriging method described could be applied directly following a training cycle to estimate system responses, a challenge with surrogate modeling is that different training processes or model initializations will produce different surrogate models of differing accuracy. One approach to combat this variability is the method of Genetic Aggregation, which leverages a universal criterion applicable to any particular type of surrogate model to support the automatic selection of an optimal ensemble of surrogate models. Figure 2.2 summarizes the Genetic Aggregation method used herein, for which details are provided in the following discussion.

Chapter 5.0 of this work employs this Genetic Aggregation algorithm which aims to construct an optimal aggregate surrogate model (Ben Salem & Tomaso, 2018), and is denoted with abbreviation GA. Thus, the resultant surrogate model is not one mathematical model (e.g., Kriging or polynomial regression) but is rather an ensemble of N_M surrogate models $\mathcal{M}^*$, with an estimation produced through a weighted combination of those models as:

$$\mathcal{M}^{GA}(\boldsymbol{x}) = \sum_{j=1}^{N_M} w_j \, \mathcal{M}_j^*(\boldsymbol{x}) \tag{42}$$

Details on the sub-surrogates with which these aggregate surrogates are constructed may be found in (Lancaster & Salkauskas, 1981) for moving least squares, (Smola & Scholkopf, 1998) for support vector regression, and (Abrahamsen, 1997) for various correlation functions that may be used for Kriging.

Let the computational data ("design points" or "training points") be established from N realizations of a full-order finite element model providing $\boldsymbol{y}$ given inputs $\boldsymbol{x}$ thus forming the set of design points $\mathbf{z}= (\boldsymbol{x}, \boldsymbol{y})$.

As discussed in (Ben Salem & Tomaso, 2018), the weight factors are determined by minimizing the error from multiple perspectives. Specifically, a penalized predictive score (PPS) is assigned to each candidate surrogate with respect to the computational data set $\boldsymbol{z}$ which combines three terms: internal accuracy measured as a root mean squared error (RMSE) $\mathcal{R}_{RMSE}$, predictive capability measured by k-fold cross validation $\mathcal{R}_{k-CV}$, and a roughness penalty $\mathcal{R}_{\mu}$.

$$PPS(\mathcal{M}^*, \boldsymbol{z}) = \Lambda_{\alpha}\mathcal{R}_{RMSE} + \Lambda_{\beta}\mathcal{R}_{k-CV} + \Lambda_{\gamma}\mathcal{R}_{\mu} \tag{43}$$

where the coefficients weight the contribution of each term for the PSS and, in this implementation, are related as $\Lambda_{\alpha} = 2\Lambda_{\beta}$ and $\Lambda_{\beta} = 2\Lambda_{\gamma}$. The formulation for the $\mathcal{R}_{RMSE}$ and $\mathcal{R}_{k-CV}$ terms are detailed in Section 2.2 of (Ben Salem & Tomaso, 2018), and $\mathcal{R}_{\mu}$ is based on a Bending Energy Functional per (Duchon, 1977).

Having thus established some universal metric by which the accuracy of each model may be measured, the question then becomes how to best weight each surrogate model (i.e., find w_j for $\mathcal{M}_j^*$ from Equation (42)). Per (Ben Salem & Tomaso, 2018), the weight factor values are based on a correlation matrix of the error terms used in establishing the PPS under the constraint that $\sum_j^{N_M} w_j = 1$ as:

$$w = \frac{C^{-1}[I]}{[I]^T C^{-1}[I]} \tag{44}$$

in which C is the matrix of the PPS errors E_i and E_j (i.e., $E_j = PPS\left(\mathcal{M}_j^*\right)$) of candidate surrogates multiplied by one another:

$$C_{ij} = \frac{1}{N} E_i E_j \tag{45}$$

Recognizing that there is a vast combination of plausible surrogate models (e.g., different mean and correlation kernels of a Kriging model), a genetic algorithm is used to increase the probability of achieving the most effective aggregate surrogate model by solving the optimization problem:

$$\min_{W} PPS(\mathcal{M}^{GA}, \boldsymbol{z}) \tag{46}$$

Thus, this is why the surrogate modeling method is termed "Genetic Aggregation"; it uses a genetic algorithm by which sequential generations determine the optimal aggregate surrogate model. The genetic algorithm is a well-accepted technique for stochastic optimization, and the implementation found within ANSYS DesignXplorer (ANSYS, 2016) was used for this work.

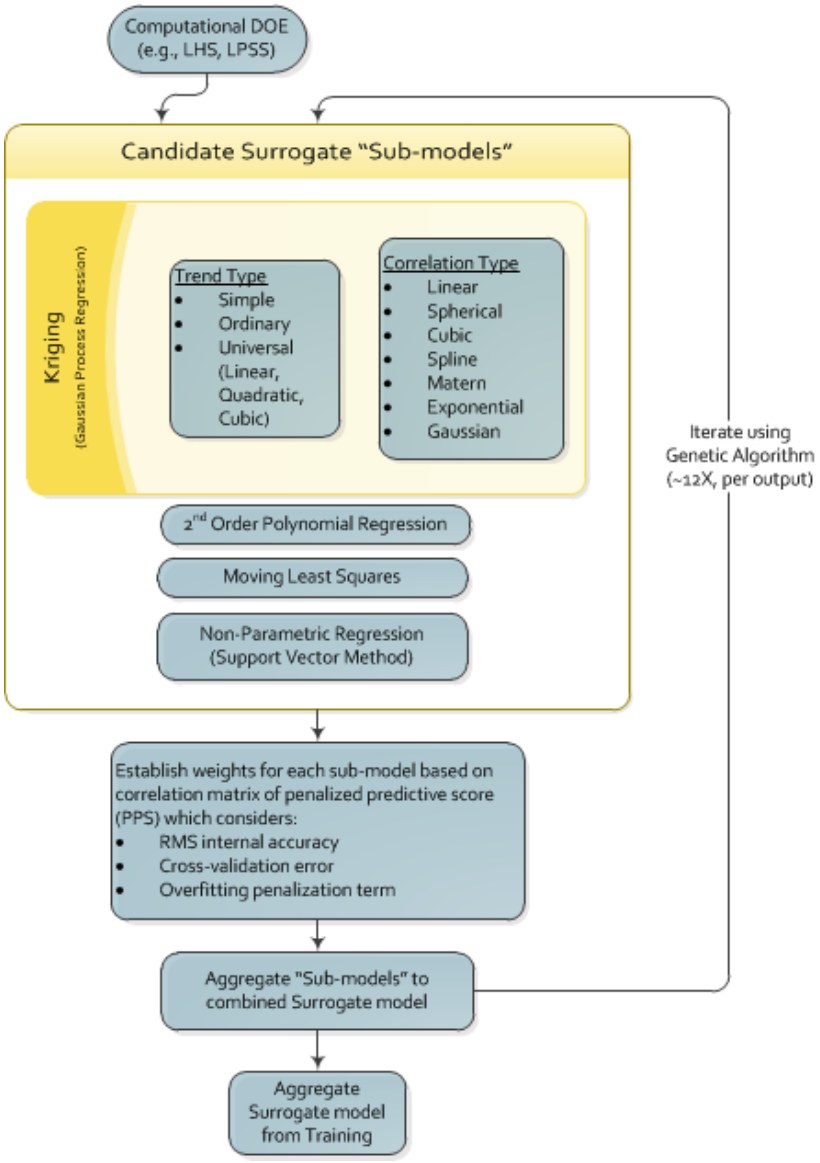

Figure 2.2 Flow Chart of Genetic Aggregation Method

2.3.3 Verification Criteria

Given the following nomenclature, Table 2.2 shows various goodness-of-fit criteria with which surrogate model quality may be assessed, as a means of verification. Within this work MRR and RRMS metrics are used, as they offer slightly different perspectives on surrogate model accuracy. The MRR provides a "worst case" error metric pertaining to the surrogate solution which is most different from the full-order model solution, whereas the RRMS provides an averaged measure of error across all N design points.

These verification criteria are applied to the Kriging surrogates (per Section 2.3.1) used in Chapters 3.0 and 4.0 , as well as to the Genetic Aggregation surrogates (per Section 2.3.2) used in Chapter 5.0 .

y_i	Value of the output parameter at the i^{th} sampling point
$\hat{y}_i$	Value of the Surrogate model $\mathcal{M}$ at the i^{th} sampling point
$\bar{y}$	Arithmetic mean of the values of y_i
σ_y	Standard deviation of the values of y_i
N	Number of sampling points (i.e., "design points")

Table 2.2 Surrogate Model Goodness of Fit Criteria

Criterion	Equation
Coefficient of Determination (R^2)	$1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}$
Maximum Relative Residual (MRR)	$\max_{i=1:N}\left(\left\lvert\frac{y_i - \hat{y}_i}{y_i}\right\rvert\right)$
Root Mean Square Error (RMSE)	$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}$
Relative Root Mean Square Error (RRMS)	$\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{y_i - \hat{y}_i}{y_i}\right)^2}$
Relative Maximum Absolute Error	$\frac{1}{\sigma_y}\max_{i=1:N}(\lvert y_i - \hat{y}_i\rvert)$
Relative Average Absolute Error	$\frac{1}{\sigma_y}\frac{1}{N}\sum_{i=1}^{N}\lvert y_i - \hat{y}_i\rvert$

2.4 Design of Experiment Sampling Strategies

A critical component to creating any surrogate model, including Gaussian Process Regression, as detailed in Section 2.3, is the training dataset. The surrogate model is only as good as the training dataset, and therefore, the dataset should be ensured to contain the important features of the system of interest. However, it is also important to attempt to minimize the size of the training dataset required, as generating this data is typically the dominating computational expense for any associated problem. Different variants of stratified sampling are studied herein. In stratified sampling, the sample space for an input parameter is divided into strata, and input values are obtained by sampling separately from within each stratum instead of from the distribution as a whole, as is done for random Monte Carlo sampling. It has been shown that stratified sampling, such as LHS, can yield substantially improved sampling errors in terms of properly characterizing a probability density function with a given number of samples (Chrisman, 2014). In fact, for univariate analysis it has been shown that the sampling error of Monte Carlo goes down as the order of $\frac{1}{\sqrt{N}}$, whereas the sampling error for LHS decreases as the order of $\frac{1}{N}$ per (Aistleitner, et al., 2012) and (Loh, 1995), for example.

Specifically, the training data sets were generated by exercising a full-order finite element model for which computational DOEs were constructed using Latin Hypercube Sampling (LHS) and a generalized Latin Hypercube sampling method called Latinized Partially Stratified Sampling (LPSS) (Shields & Zhang, 2016).

2.4.1 Latin Hypercube Sampling

In contrast to random Monte Carlo sampling, LHS is a way to randomly sample a parameter space that aims to spread the sample points evenly across all possible values (McKay, et al., 1979). One of its first implementations of LHS was in a Sandia National Laboratory computer program simply entitled "Latin Hypercube Sampling" (Iman, et al., 1980). As a version of stratified sampling, LHS partitions each input parameter distribution into intervals of equal probability, selects one sample from each interval, and shuffles the sample for each input so that there is no correlation. To generate M samples using LHS, each input distribution is divided into M intervals of equal probability. A given sample is generated by selecting one value at random from each of the inputs, but without replacement, from the M sample values for each input. This results in M samples, with each input being used only once (Morgan, et al., 1992).

LHS operates by dividing each vector component (parameter) p_i into M disjoint subsets (strata) of equal probability Ξ_{ik}; where $i = 1,2,\dots,N$ and $k = 1,2,\dots,\mathrm{M}$. Samples of each parameter are drawn from the respective strata according to Equation (47).

$$p_{ik} = D_{X_i}^{-1}(U_{ik}) \tag{47}$$

where U_{ik} are independent, identically, and uniformly distributed samples on the domain $\left[\xi_k^l, \xi_k^u\right]$ with $\xi_k^l = \frac{k-1}{M}$ and $\xi_k^u = \frac{k}{M}$. A term p_{ik} is randomly selected from each parameter (without replacement) and these terms are grouped to produce a sample, and then that process is repeated M times (Shields & Zhang, 2016).

2.4.2 Latinized Partially Stratified Sampling

The LPSS method developed in (Shields & Zhang, 2016) performs simultaneous Latin sampling of all variables and stratified sampling of subsets of variables, and has been shown to provide variance reduction in the context of parameter interactions. In the case of simulating FIV phenomena and system finite element modeling, such as explored within this dissertation, often parameter interactions can exist in terms of (uncorrelated) forcing functions or physical model parameters.

LPSS is achieved by first defining a partially stratified sampling (PSS) design, as described in (Shields & Zhang, 2016), and as follows. Let $\Theta_i, i = 1, \dots, N_S$ denote N_S disjoint N_i-dimensional orthogonal subspaces of the N-dimensional sample space. PSS divides each subspace Θ_i into a collection of M_i disjoint subsets Ξ_{ik}; $k = 1,2, \dots, M_i$. Lower N_i dimensional random samples $\boldsymbol{p}_{ik} = \{p_{ik1}, p_{ik2}, \dots, p_{ikN_i}\}$ are generated within each stratum Ξ_{ik} of subspace Θ_i according to the stratified sampling method. Then, full N-dimensional samples $\boldsymbol{p}$ are assembly by randomly grouping the lower-dimensional samples generated in each subspace. The most significant challenge to PSS is to identify the optimal subspace decomposition. In some cases, it may be clear which variables are interacting which will inform the PSS subspace definitions, but such cases are the exception. Indeed, the variables which are interacting are not known *a priori* in the present work applied to nuclear reactor internals structural dynamics. Coupling PSS with "Latinized" stratified sampling (LSS) helps to alleviate this concern.

LSS permits one to simultaneously reduce variance associated with both the main and interaction effects by constructing, on a given N_i-dimensional subspace, a true stratified sampling design that is at the same time an LHS design. The procedure for accomplishing LSS

is described in Table 2.3. Note that this method is colloquially called "Su-Do-Ku" sampling as a low dimensional LSS grid resembles the popular game.

Table 2.3 LSS Method Procedure

Step	Description
1	Draw an LHS from the subspace
2	Stratify the domain as desired ensuring that the stratification is consistent with an LHS design
3	For each stratum, randomly select a point p_i from each component of the LHS (without replacement) such that the sample $p = \{p_1, p_2, \dots, p_N\}$ lies within the stratum
4	Repeat for each stratum of the design.

Finally, LPSS involves the use of LSS in combination with PSS. Following the definition of a PSS design, samples from the lower-dimensional subspaces are then drawn according to the LSS method. Thus, with LPSS, it is sufficient to stratify a set of variables together simply based on the possibility that they may interact. If the chosen parameters do interact then the savings, in terms of variance reduction (and associated accuracy of surrogate model and GSA, in this application), will be amplified by reduction in both the main effects and interactions. If the chosen parameters do not interact, there will be no increase in variance since the main effects are also being filtered thanks to LSS.

Implementation of the LPSS method in the form of Matlab code is shown in Appendix A.

2.4.3 Adaptive Sampling

Sampling, training, and verifying a Surrogate model in a serial manner can prove computationally inefficient if it turns out that the number of samples initially chosen to perform the computational DOE exceeds the number of (intelligently-drawn) samples actually required to achieve some desired level of model accuracy. This principle is illustrated in Figure 2.3. Let N samples fed through a computational DOE such as LHS or LPSS yield some verification error, and set that error equal to the verification tolerance. Next, let some smaller number of samples n_1 be fed through a computational DOE, such that $n_1 \ll N$, recognizing that it is likely that the verification error upon drawing n_1 samples will exceed the verification tolerance. Adaptive sampling explores the situation which, upon subsequent iterations i of adding n_R refinement points it may be that n_i samples yields equivalent verification error to that produced by N samples even while $n_i < N$ because the refinement points are intelligently drawn from places within the parameter space which best serve to reduce error.

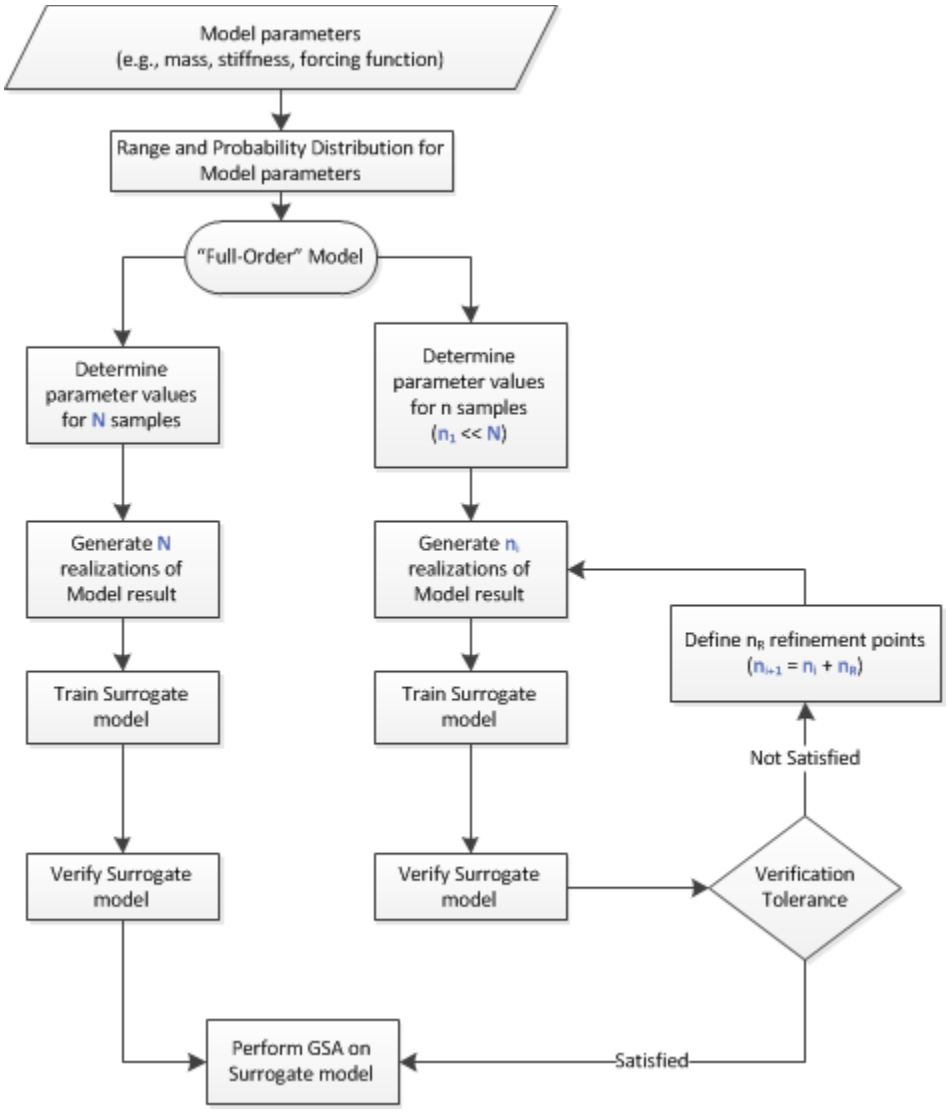

Figure 2.3 Illustration of Adaptive Sampling Principle

For the present study, the initial computational DOE used to train the surrogate model for SFEM (see Section 5.2) used LHS to randomly sample the parameter space. The LHS DOE provided a set of data with which the surrogate model was initially trained. Given the relatively high dimensionality of the SFEM sensitivity analysis, it is expected that some parameters may converge much more easily than others (i.e., require fewer samples to be adequately resolved).

Adaptive sampling provides a means by which subsequent full-order model solutions may prove most effective in achieving an accurate surrogate model so that computational expense may be minimized.

"Refinement Points" were thus employed when constructing and verifying the surrogate model. Adaptive sampling was accomplished for the present study by the Universal Prediction-based Surrogate Modeling Adaptive Refinement Technique (UP-SMART) algorithm described in (Ben Salem, et al., 2017). The following briefly describes the UP-SMART algorithm.

Let $\mathcal{M}$ denote a real-valued function defined such that $\boldsymbol{y} = \mathcal{M}(\boldsymbol{x})$. In the present application, $\mathcal{M}$ represents the finite element model. $\mathcal{M}$ can be estimated by obtaining a set of N samples which relate the inputs $\boldsymbol{x}$ to the outputs $\boldsymbol{y}$; let $\boldsymbol{z}$ denote these observations. These observations are the "design points" solved, as determined by the DOE sampling. Using the observations $\boldsymbol{z}$ a Kriging surrogate model $\mathcal{M}^K$ is constructed which mimics the behavior of $\mathcal{M}$ per methods described in Section 2.3.1. For exploration of adaptive sampling in this dissertation, an aggregate surrogate model $\mathcal{M}^{GA}$ is used, as described in Section 2.3.2.

The UP-SMART adaptive sampling method draws points based on the cross-validation error of the aggregate surrogate model. The idea of surrogate model verification recognizes that training and evaluating statistical performance of a surrogate model on the same data with which it was trained yields an optimistic (and thus misleading) result (i.e., the surrogate model may not generalize to accurately estimate unseen samples). Typically, at least one portion of the data is assigned as the training set while the remainder is used for testing. For more thorough analysis, cross-validation splits the data multiple times and evaluates the predictive capabilities of a model when trained with and tested against multiple splits of the same data set. For example, the "leave-one-out" cross-validation methods leaves one point out of the population of data and uses

the remainder of the data set to train the model to predict that (excluded) point. The predictions of the left-out points provide a distribution across the parameter space of which locations can be best (or worst) predicted. It is upon such a distribution, termed the Universal Prediction (UP) distribution, that new samples are defined with the adaptive sampling method used herein.

The UP distribution (Ben Salem, et al., 2017) is defined as a weighted empirical distribution, in which the weights κ are computed as:

$$\kappa_{i,n}(x) = \frac{1 - e^{-\frac{d\left((x,x_i)\right)^2}{\rho^2}}}{\sum_{j=1}^{N}\left(1 - e^{-\frac{d\left((x,x_j)\right)^2}{\rho^2}}\right)} \tag{48}$$

where $d\big((x, x_i)\big)$ is the distance between the i^{th} design point and the input parameter value x, and ρ is a smoothing parameter. Thus, $\kappa_{i,n}(x)$ increases with the distance between the i^{th} design point x_i and x. Using this smoothed weight function, the UP distribution is defined as:

$$\mu_{(n,x)}(dy) = \sum_{i=1}^{n} \kappa_{i,n}(x)\delta_{\mathcal{M}^{GA}_{n,-i(x)}}(dy) \tag{49}$$

(Ben Salem, et al., 2017) states that this probability measure is "the empirical distribution of all the predictions provided by cross-validation sub-models weighted by local smoothed masses". Thus, the Dirac function $\delta_{\mathcal{M}^{GA}_{n,-i(x)}}$ and the local smoothed mass term is $\kappa_{i,n}(x)$ provide this empirical distribution.

The local UP variance and mean are then defined within (Ben Salem, et al., 2017) as:

$$\widehat{m}_n(x) = \int y\mu_{(n,x)}\,(dy) = \sum_{i=1}^{n} \kappa_{i,n}(x)\mathcal{M}^{GA}_{n,-i(x)}(x) \tag{50}$$

$$\hat{\sigma}_n^2(x) = \int \left(y - \hat{m}_n(x)\right)^2 \mu_{(n,x)}\,(dy)$$
$$= \sum_{i=1}^{n} \kappa_{i,n}(x) \left(\mathcal{M}^{GA}_{n,-i(x)}(x) - \hat{m}_n(x)\right)^2 \tag{51}$$

The UP-SMART thus samples points where the UP distribution variance, as defined by Equation (51), is maximal. During the refinement iterations, selection of the parameter value for the $n+1$ point is determined as:

$$x_{n+1} \in \underset{x \in \mathbf{X}}{\operatorname{argmax}}\left(\gamma_n(x)\right) \tag{52}$$

$$\gamma_n(x) = \hat{\sigma}_n^2(x) + \delta \underline{d}_{X_n}(x) \tag{53}$$

Where:

$\hat{\sigma}_n^2(x)$	UP variance at point x for step n
$\underline{d}_{X_n}(x)$	A given distance on $\mathbb{R}^p$
X_n	The sample set used to build the surrogate model
$\mathbf{X}$	User-defined parameter space
δ	Distance penalization

In recent years adaptive sampling techniques have found some interest in applications related to nuclear energy, such as in (Idaho National Laboratory, 2015), but have not gained widespread acceptance or use. It is therefore the intent of this dissertation to provide an introduction as to the potential value offered by adaptive sampling, by implementing one such method (which is not necessarily claimed to be optimal but rather illustrative), as discussed in 5.0 .

2.5 Integrated Computational Framework

In the interest of summarizing how the methods described in Chapter 2.0 fit together within this work, Figure 2.4 illustrates how a random vibration finite element analysis (Section 2.1) may be used to accomplish a GSA (Section 2.2) by using a Surrogate model (Section 2.3) trained by a computational DOE (Section 2.4), thus forming a computational framework with which experiments are performed in Chapters 3.0 for an UI FIV problem and 4.0 for the SFEM. The novelty of this work lies not necessarily in any one of the particular methods described in the preceding subsections, but rather in the integration of these methods in such a workflow for a nuclear reactor application.

Beginning with a set of data determined via a computational DOE, such as LHS or LPSS, Kriging surrogate models were trained with which GSA was performed. In parallel to using surrogates, GSA was performed directly on the full-order finite element models in order to establish benchmark sensitivities. These sensitivities were compared in order to study the error associated with surrogate-based GSA and the role of the sampling method with which the surrogates were trained. This is studied using two separate finite element models; one a stationary FIV of a reactor UI and the other a SFEM subjected to LOCA loading.

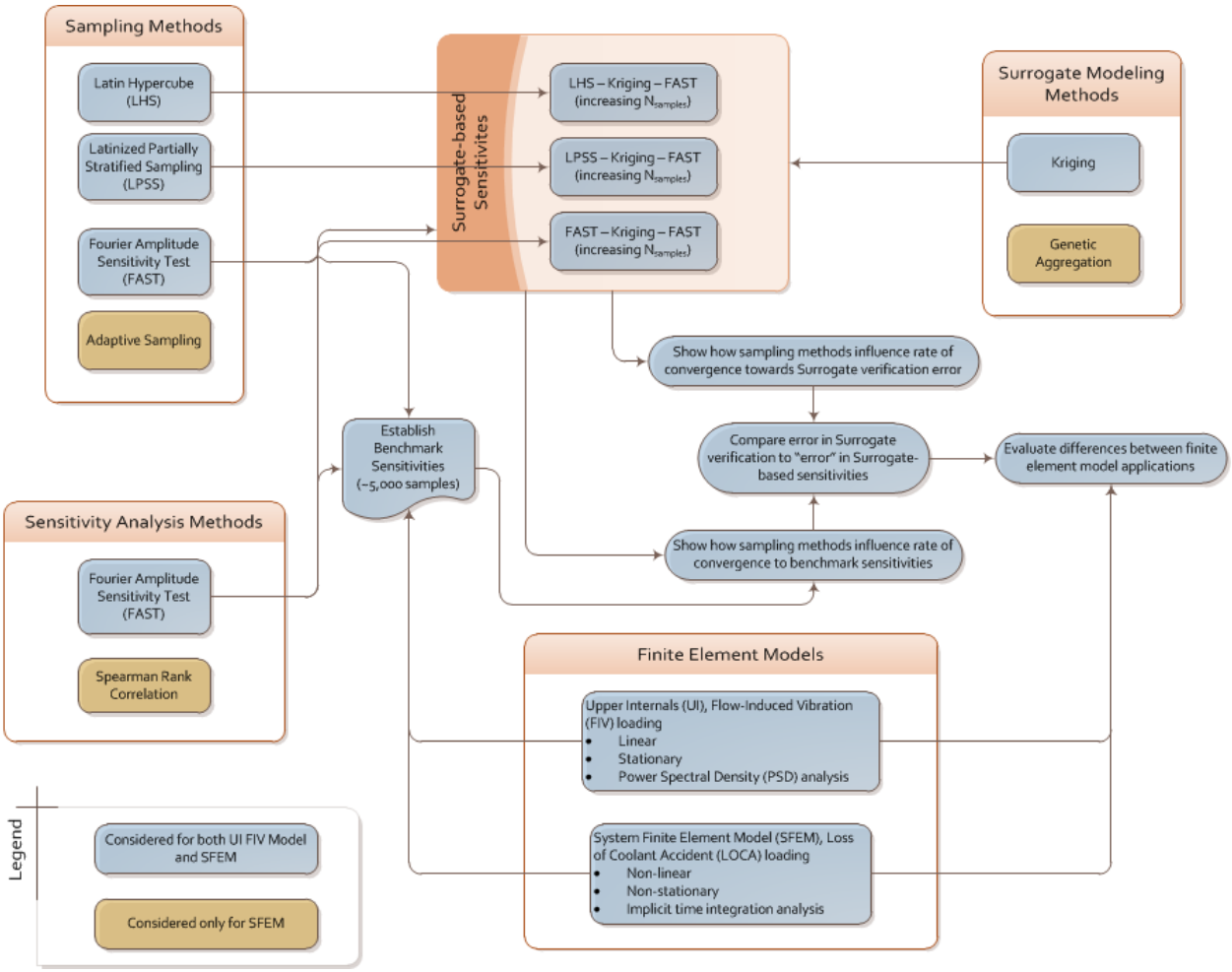

Figure 2.4 Integrated Workflows

3.0 Application to Stationary Random Vibration for Upper Internals Assembly

According to the introduction to stochastic dynamics for nuclear reactor structures described in Chapter 1.0 , and employing the methods described in Chapter 2.0 , the response of the Upper Internals (UI) assembly of a nuclear reactor subject to FIV is investigated in this Chapter (Banyay, et al., 2019). Particularly within the upper plenum of a reactor the flow is chaotic. Decomposition of the various flow characteristics into distinct forcing functions (e.g., jetting, cross-flow, etc.) constitutes a challenge unto itself, which is not specifically addressed by this dissertation (see (Au-Yang, 2001) for example). That said, once distinct forcing functions have been established, a finite element analysis is used to simulate the structural dynamics and thus compute displacements, strain, and stress. The structures themselves are generally constructed of geometrically simple shapes, in part, due to ASME Code design guidelines, as shown in Subsection NG of (ASME, 2017), and also for purposes of manufacturability and regulatory acceptance. As such, various structures, such as lower and upper support columns, can be rightly approximated as a combination of axisymmetric beams, as well as shell elements and lumped masses within a finite element model. Thus, while the computational model is intended to simulate a complex system, simplicity of some features in the mechanical component design permit the finite element analysis to credibly use such element types as beams and shells. As an example, Figure 3.1 shows an UI assembly for the AP1000 plant.

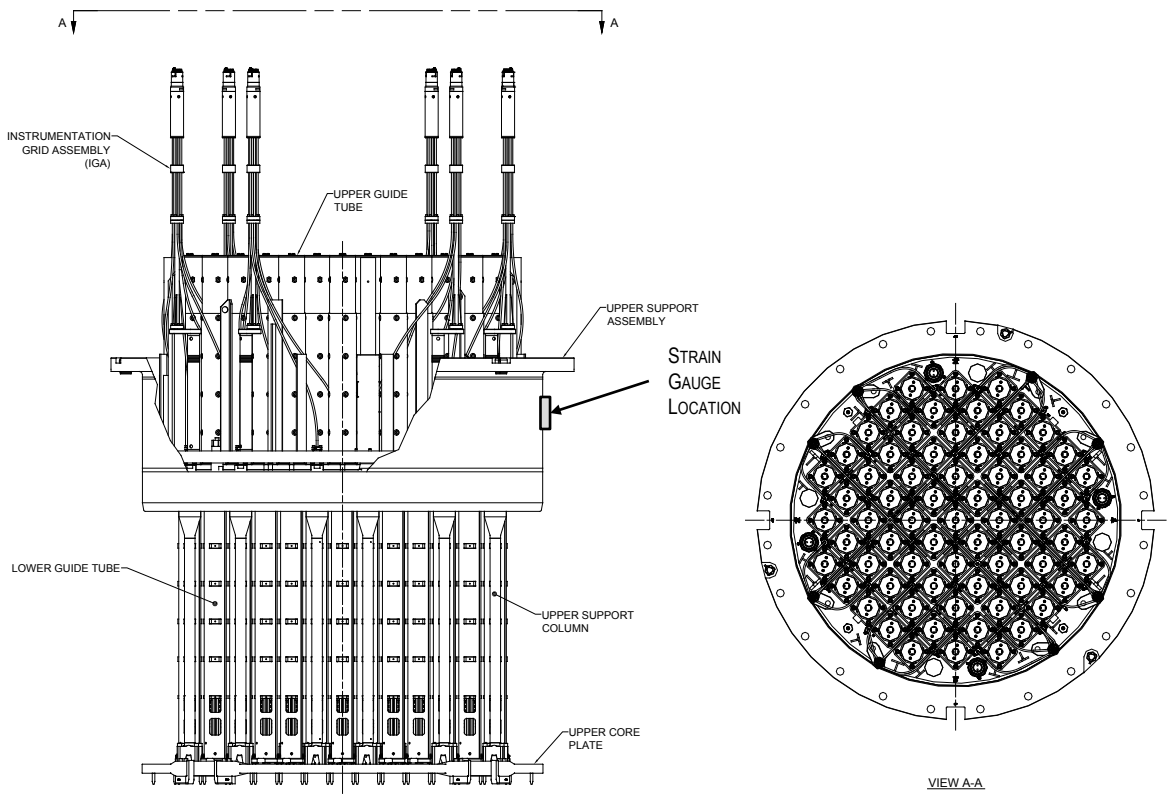

Figure 3.1 Upper Internals Assembly for the AP1000 Plant (from (Westinghouse Electric Company, 2011)

3.1 Finite Element Analysis Description

A finite element model of the UI section of the reactor assembly was built in ANSYS using linear hexahedral elements. All structural components were taken to be austenitic stainless steel with density 7,850 $\frac{kg}{m^3}$, Poisson's ratio 0.3, and elastic modulus $2 \times 10^{11} Pa$. The material was considered linear elastic and geometric nonlinearities were not considered, as small displacement theory is considered a reasonable assumption for FIV during normal operating conditions of a reactor. The structure was supported with simple supports at the top rim of the upper support skirt, in the sense that displacement was zero in all three translational degrees of freedom, and subjected to three direct-applied forcing functions to the support columns acting in the radially outward direction, lateral forcing functions on each of the upper core plate and upper support plate acting in mutually perpendicular directions, and a vertical forcing function on the upper core plate. The model, including boundary conditions and loads is shown in Figure 3.2. The output of interest from this model was assumed to be the axial (normal) strain, denoted ϵ, of the upper support skirt in the interest of simulating a virtual strain gauge measurement; see Figure 3.2.

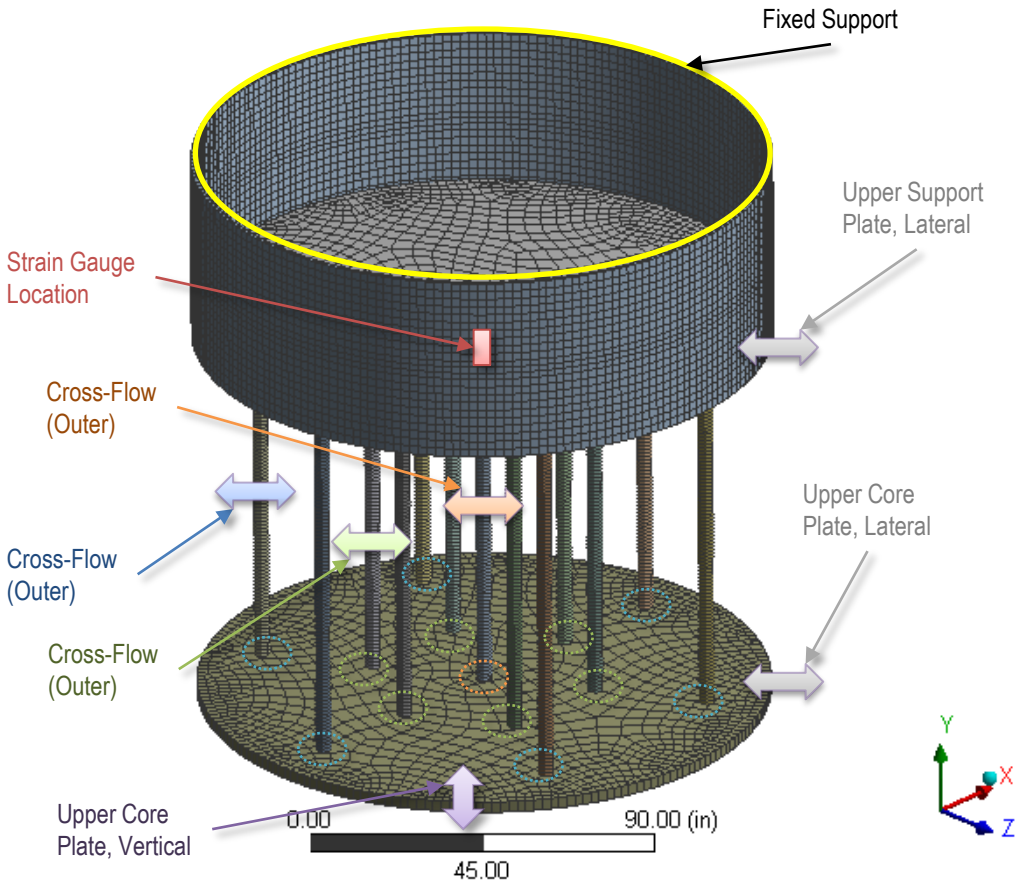

Figure 3.2 Upper Internals Model, Mesh, and Boundary Conditions

Two types of loads were applied to the structure:

1. Cross-flow loads on the support columns, and
2. Base motions applied to the upper core plate and upper support plate.

For the FIV problem, all forcing functions are defined as force or acceleration PSDs acting on un-constrained nodes (i.e., not imposed at the support locations). Cross-flow loads correspond to forces induced on the columns by turbulent flow of coolant over the columns. Here, cross-flow loads were modeled as a stochastic process that is fully correlated along the length of the

column. To more accurately predict forced response, the cross-flow loading could be defined with an uncertain correlation length and permitted to vary between one and three diameters along the length of the column (consistent with that observed by Mulcahy for turbulent cross-flow (Mulcahy, 1982)). However, this length variation was considered a secondary effect and, for this study, was not considered. The term "base motion" is used for displacements applied to the upper core plate and upper support plate as those loads are caused by adjacent reactor components which were not included in the finite element model (e.g., base motion imparted from the core barrel to the upper support plate).

Both the direct-applied and base motion loading were applied from 0 – 2,000 Hz with exponential spectral decay of $G(\omega) \propto \omega^{-1.75}$, as illustrated in Figure 3.3. This spectral decay appears consistent with the non-dimensional forced response PSDs provided by (Au-Yang & Jordan, 1980) and (Mulcahy, 1982). For comparison, forced response data from a column-like structure exposed to cross-flow loading from a CVAP hot functional test program is shown in Figure 3.3 with dashed lines. The peaks in the forced response correspond to the natural frequencies of the structure, but it may be seen that the broadband decay of the applied forcing function is consistent with the data. Statistical analysis of this same dataset reveals that the forced response is Gaussian, ergodic and stationary.

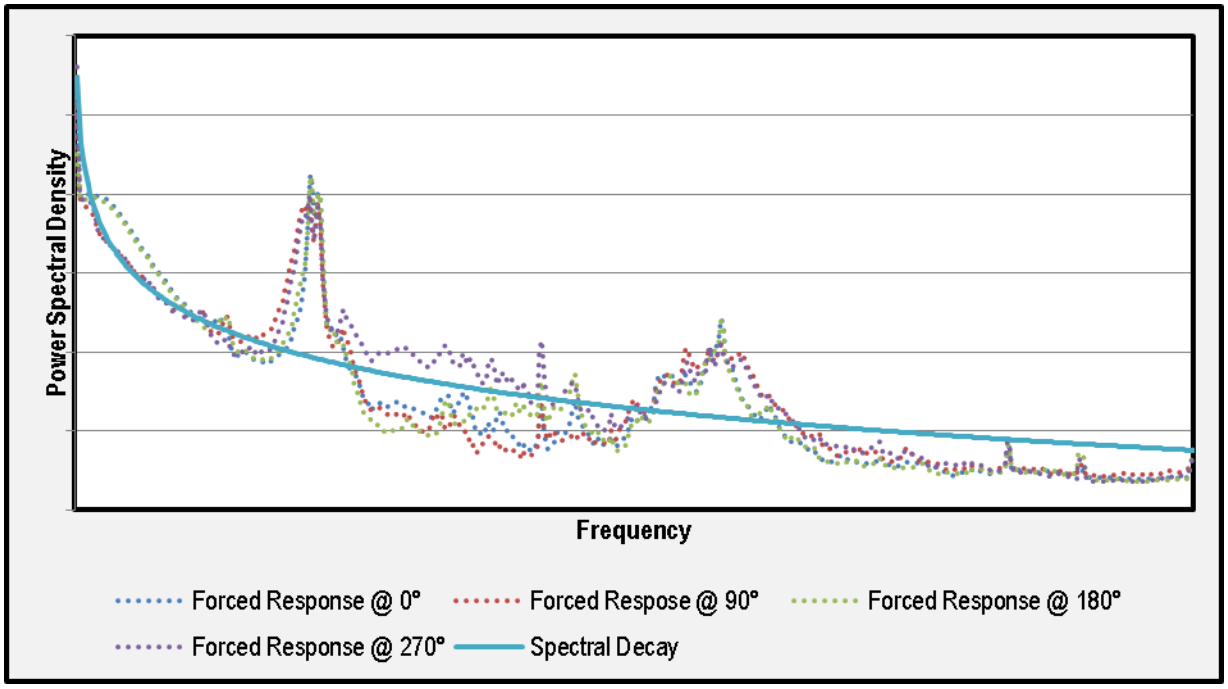

Figure 3.3 Forcing Function Non-Dimensional Power Spectral Density (with beam forced response data overlaid)

For GSA, the magnitude of the PSD was scaled as shown by the λ terms in Table 3.1. In this analysis, the λ terms were assigned a Uniform distribution on the range $[0.9, 1.1]$, from which random samples were drawn for the computational DOE. The magnitude of the force and displacement PSDs were chosen based on the approximate order of magnitude of which these loads have been recorded from various historical instrumented Hot Functional Tests and sub-scale tests of PWRs.

Table 3.1 Forcing Functions Applied to Upper Internals Finite Element Model

Forcing Function	Type of Excitation	PSD Scaling
Cross-Flow across Support Columns	Force	$\tilde{G}_x^F(\omega)_{USC.Center} = \lambda_{USC.Center} \times G_x^F(\omega)$ $\tilde{G}_x^F(\omega)_{USC.Mid} = \lambda_{USC.Mid} \times G_x^F(\omega)$ $\tilde{G}_x^F(\omega)_{USC.Outer} = \lambda_{USC.Outer} \times G_x^F(\omega)$
Base motion applied to Upper Core Plate and Upper Support Plate	Displacement	$\tilde{G}_x^\delta(\omega)_{UCP.Edge} = \lambda_{UCP.Edge} \times G_x^\delta(\omega)$ $\tilde{G}_z^\delta(\omega)_{USP.Edge} = \lambda_{USP.Edge} \times G_z^\delta(\omega)$ $\tilde{G}_y^\delta(\omega)_{UCP.Face} = \lambda_{UCP.Face} \times G_y^\delta(\omega)$

3.2 Sensitivity Analysis of Stationary Random Vibration for UI Assembly

3.2.1 Sensitivity Analysis of Full-Order Model using FAST

The FAST method (Section 2.2.1) was applied directly using the finite element analysis (i.e., full-order model) in order to compute benchmark global sensitivities against which those obtained from various Surrogate models could be compared. Using the SAFE toolbox documented in (Pianosi, et al., 2015) (for which the underlying methodology is similar to that implemented in UQLab (Marelli & Sudret, 2014)), GSA was performed. An increasing number of samples were generated for the 6 inputs to the full-order model until convergence of the sensitivity indices was observed. The number of simulations required to produce sensitivity indices that did not change more than 0.3% (or a sensitivity index magnitude of 0.0005) upon further samples was 5,000. Convergence of the sensitivity indices is plotted in Figure 3.4, from

which it may be seen that there was no substantial change in the magnitude of the sensitivity indices as sample sizes greater than 512. Specifically, between 1,728 and 5,000 samples, the sensitivity indices associated with the base motion loads differed by less than 0.0005 and those associated with cross-flow loads differed by less than 3.6×10^{-5}.

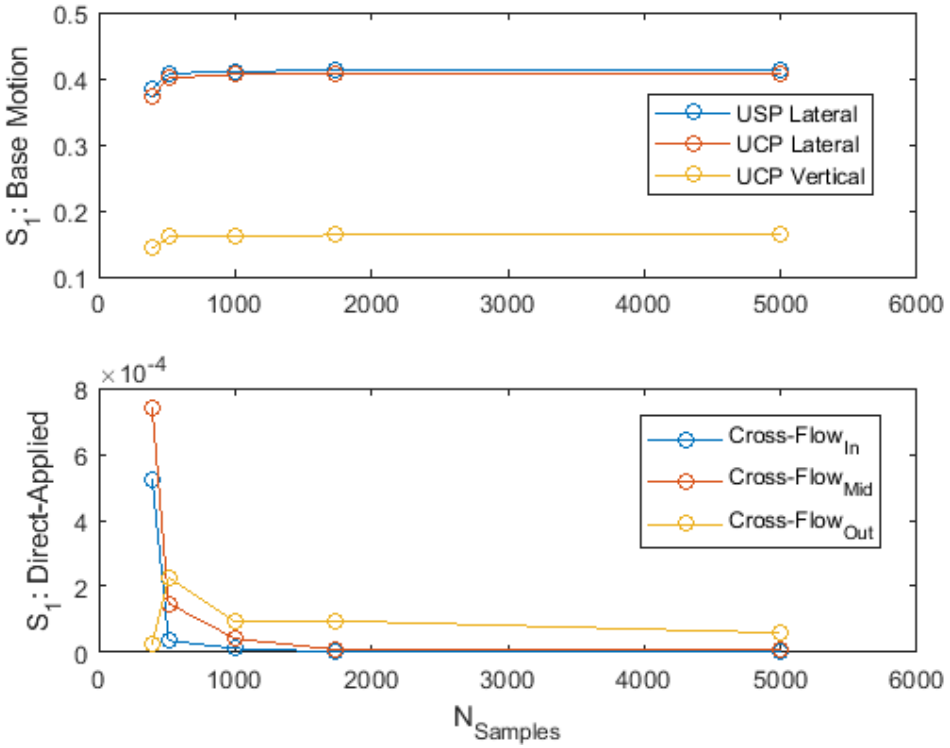

Figure 3.4 Full-order Model GSA Convergence

Prior to examining results using surrogate modeling in place of the full-order model, the physical significance of the GSA results is noteworthy. The GSA results show that the base motion loads acting in the lateral direction at both the UCP and USP each account for 41% of the output variance, the UCP vertical forcing function accounts for 16% of the output variance, and the sum total of the direct-applied cross-flow loads account for less than 1% of the output

variance. This is meaningful in the sense that the dynamic response of the UI assembly is governed more by turbulence that imparts loads to the interface joints (i.e., upper support flange) than turbulence that acts directly upon the upper support columns.

This sort of observation serves to inform key engineering decisions related to up-front design investments as well as in diagnostics during plant operation. For example, to borrow a principle of decision theory, the expected value of perfect information would be much greater to define the forces associated with base motion loads than direct-applied cross-flow loads, during the design stage of a new reactor. Correspondingly, given the extreme difficulty of placing sensors within an operating reactor environment, it is valuable to understand that measuring the flow field directly by placing a sensor within the upper plenum would provide very limited insight into the structural dynamic behavior of the UI structures.

3.2.2 Surrogate-based Sensitivity Analysis

Multiple avenues were evaluated by which a Kriging surrogate, per Section 2.3.1, is trained from various computational data sets. The global sensitivities obtained from the Surrogate models in these subsections are determined using FAST (Section 2.2.1). In particular, it is of interest to examine the change in GSA results having to do with increasing the sample size and changing the sampling method. To that end, the three workflows illustrated in Figure 2.1 were employed for the GSA.

3.2.2.1 Surrogate trained by FAST Samples

Using the model evaluations from the FAST GSA (addressed in Section 3.2.1), a Kriging surrogate was built for four sample sizes (i.e., the number of samples drawn from the full-order model used to train the surrogate). Given that this problem had 6 random inputs, the minimum number of samples drawn from the full-order model and evaluated using FAST was 393 (Cukier, et al., 1978), and the subsequent three sample sizes considered were 512, 1,000, and 1,728.

Then, in order to provide credibility of using a surrogate for GSA, sensitivity indices were determined via FAST using 10,000 samples drawn from these surrogate models (implemented in SAFE (Pianosi, et al., 2015)), which are shown in Figure 3.5, and the surrogate-based sensitivities were compared to those computed directly from the full-order model. The black circles on Figure 3.5 represent the benchmark sensitivities against which the surrogate-based sensitivities are compared (i.e., the sensitivities associated with 5,000 samples from Figure 3.4), and the differences between the surrogate-computed first order sensitivity indices S_1^{Surr} and those computed from the full-order model S_1^{FEA} are then shown on Figure 3.6.

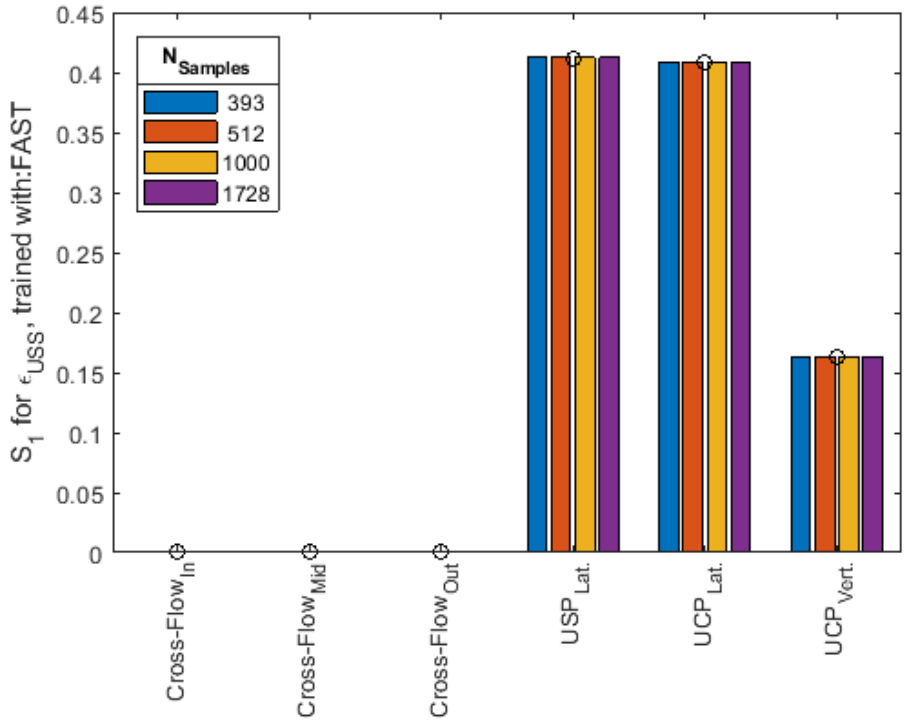

Figure 3.5 GSA from Kriging Surrogates trained with FAST at varying Sample Sizes

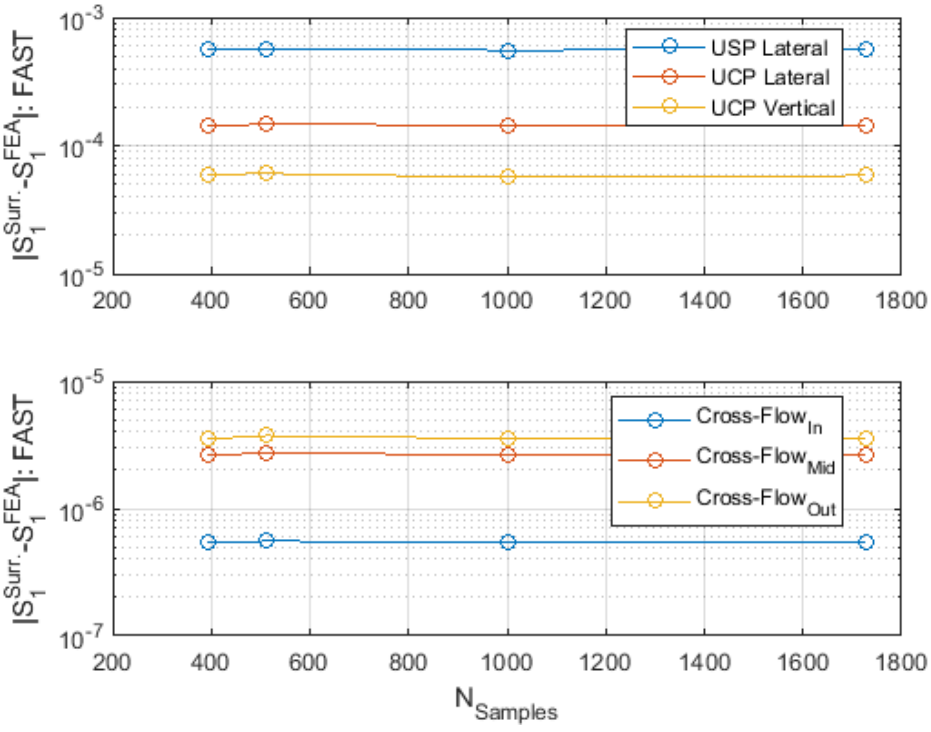

Figure 3.6 FAST-trained Surrogate-based Sensitivities vs Benchmark

3.2.2.2 Surrogate trained by LHS and LPSS

The effectiveness of using the LHS and LPSS sampling methods described in Sections 2.4.1 and 2.4.2, respectively, in establishing an accurate surrogate with which sensitivities can be calculated was evaluated. Using workflows 2-3 from Figure 2.1, the two different sampling methods of computational DOEs were used to construct a surrogate model for GSA. Following workflow 2, the Kriging surrogate was constructed from samples generated by LHS. Following workflow 3, the Kriging surrogate was built from samples generated from LPSS. The LPSS designs assumed the crossflow loads were grouped for stratification and the base motion displacement loads were grouped for stratification. This LPSS design is described in Table 3.2 and results in samples sizes of 27, 64, 125, 512, 1000, and 1728 samples from which the surrogate model was trained. LHS of the same size were used for a fair comparison.

Table 3.2 Setup of Partially Stratified Sampling Design

forcing function	**sub-domain dimension (N_i)**	**strata**	**number of samples**
cross-flow (inner, middle, outer on upper support column)	3	3 4 5 8 10 12	27 64 125 512 1000 1728
base motion (upper core plate, upper support plate, vertical)	3		
number of dimensions	6		

All global sensitivity analyses were again performed by drawing 10,000 samples from the surrogates using FAST as implemented within (Pianosi, et al., 2015). For each sample size (i.e., the number of samples used to train the surrogate), the resultant first-order Sensitivity indices are shown in Figure 3.7 and Figure 3.8 based on LHS and LPSS, respectively, in which the black

circles are the benchmark sensitivities computed from the full-order model directly. It may be seen that the sensitivities are very similar between those computed from surrogates trained from LHS or LPSS, with a maximum relative difference of 0.98% on the cross-flow load sensitivities or 0.002% on the base motion load sensitivities. This data (i.e., workflows 2 and 3 from Figure 2.1), along with the surrogate-based sensitivities established from FAST sampling (workflow 1 from Figure 2.1) is shown in Table 3-3 as well. Figure 3.9 and Figure 3.10 show the differences between the surrogate-based sensitivities and those from the full-order model. Aside from the lowest sample numbers (e.g., 27 or 64), increasing the sample size beyond 125 had an almost negligible effect on the resulting surrogate-based sensitivities.

Figure 3.11 and Figure 3.12 show the results associated with the surrogates trained from FAST sampling, LHS, and LPSS overlaid on the same graph. Of significance is that even though the surrogates used to calculate these sensitivities were constructed from different sampling strategies, the resultant global sensitivities do not appreciably differ from one another, which serves to demonstrate the stability of the GSA results for this type of system.

Table 3.3 Sensitivity Indices Computed from Surrogate Models

Forcing Function	$N_{FullOrderRuns}$	Cross-Flow (Inner)	Cross-Flow (Mid)	Cross-Flow (Outer)	Upper Support Plate Lateral	Upper Core Plate Lateral	Upper Core Plate Vertical	Sum
Kriging Model built from LPSS	27	3.61×10^{-5}	6.11×10^{-5}	2.53×10^{-4}	0.4107	0.4116	0.1626	0.985
	64	3.71×10^{-7}	2.58×10^{-6}	6.31×10^{-5}	0.4124	0.4087	0.1632	0.984
	125	3.24×10^{-7}	2.92×10^{-6}	6.30×10^{-5}	0.4125	0.4086	0.1633	0.984
	512	3.16×10^{-7}	3.09×10^{-6}	6.28×10^{-5}	0.4125	0.4085	0.1634	0.984
	1000	3.17×10^{-7}	3.13×10^{-6}	6.28×10^{-5}	0.4124	0.4085	0.1634	0.984
	1728	3.16×10^{-7}	3.15×10^{-6}	6.29×10^{-5}	0.4124	0.4085	0.1634	0.984
Kriging Model built from LHS	27	8.29×10^{-6}	6.04×10^{-5}	2.87×10^{-4}	0.4270	0.3912	0.1665	0.985
	64	3.21×10^{-7}	3.14×10^{-6}	6.19×10^{-5}	0.4126	0.4085	0.1633	0.984
	125	3.18×10^{-7}	3.08×10^{-6}	6.25×10^{-5}	0.4126	0.4084	0.1634	0.984
	512	3.16×10^{-7}	3.13×10^{-6}	6.27×10^{-5}	0.4125	0.4084	0.1634	0.984
	1000	3.17×10^{-7}	3.18×10^{-6}	6.27×10^{-5}	0.4124	0.4085	0.1634	0.984
	1728	3.18×10^{-7}	3.18×10^{-6}	6.29×10^{-5}	0.4124	0.4085	0.1634	0.984
Kriging Model built from FAST GSA samples	393	3.18×10^{-7}	3.20×10^{-6}	6.28×10^{-5}	0.4124	0.4085	0.1634	0.984
	512	3.16×10^{-7}	3.14×10^{-6}	6.30×10^{-5}	0.4124	0.4085	0.1634	0.984
	1000	3.17×10^{-7}	3.19×10^{-6}	6.28×10^{-5}	0.4124	0.4085	0.1634	0.984
	1728	3.17×10^{-7}	3.18×10^{-6}	6.29×10^{-5}	0.4124	0.4085	0.1634	0.984

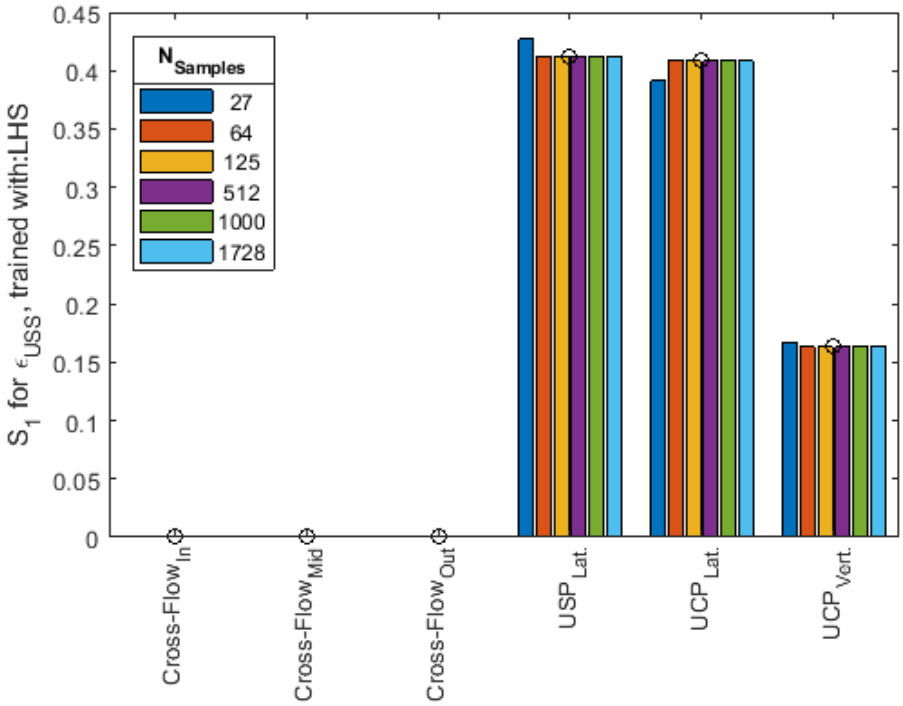

Figure 3.7 Comparison of Sensitivities from LHS-Trained Surrogate Models

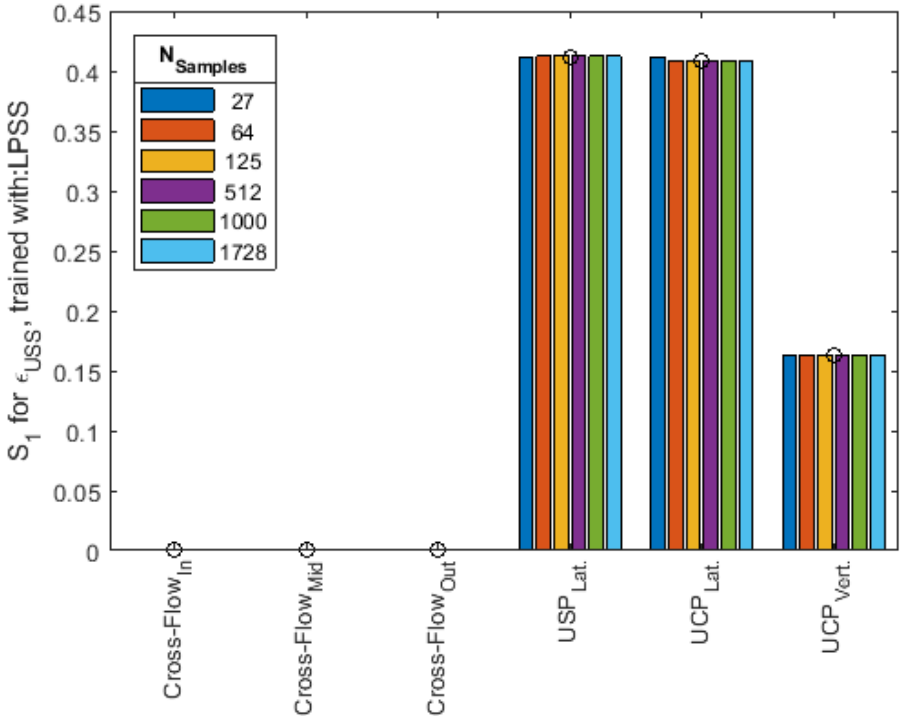

Figure 3.8 Comparison of Sensitivities from LPSS-Trained Surrogate Models

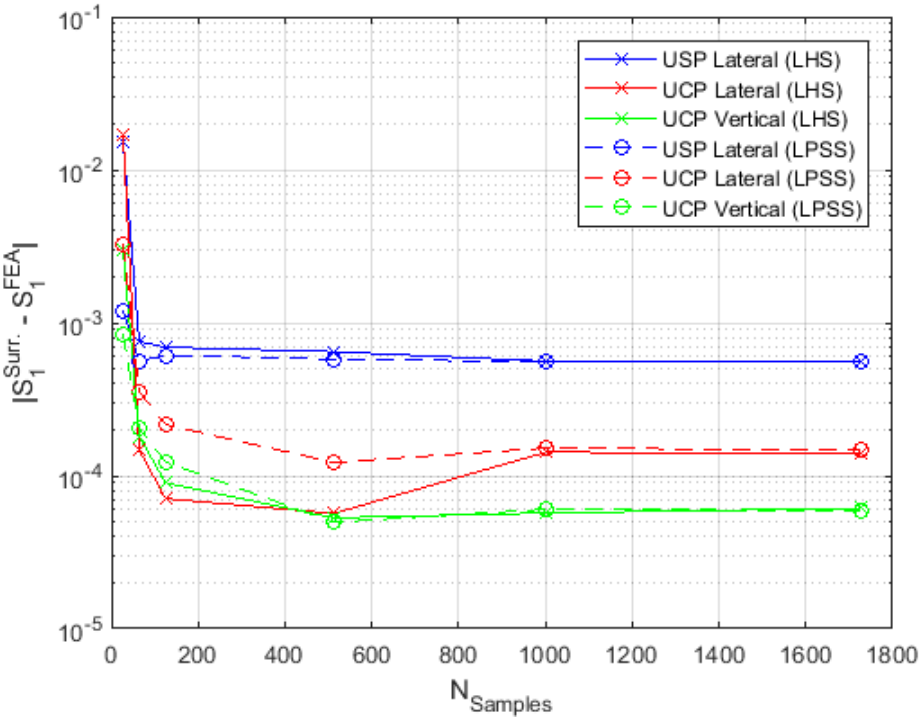

Figure 3.9 Comparison of LHS-trained and LPSS-trained Surrogate-based Sensitivities to Benchmark for Base Motion loads

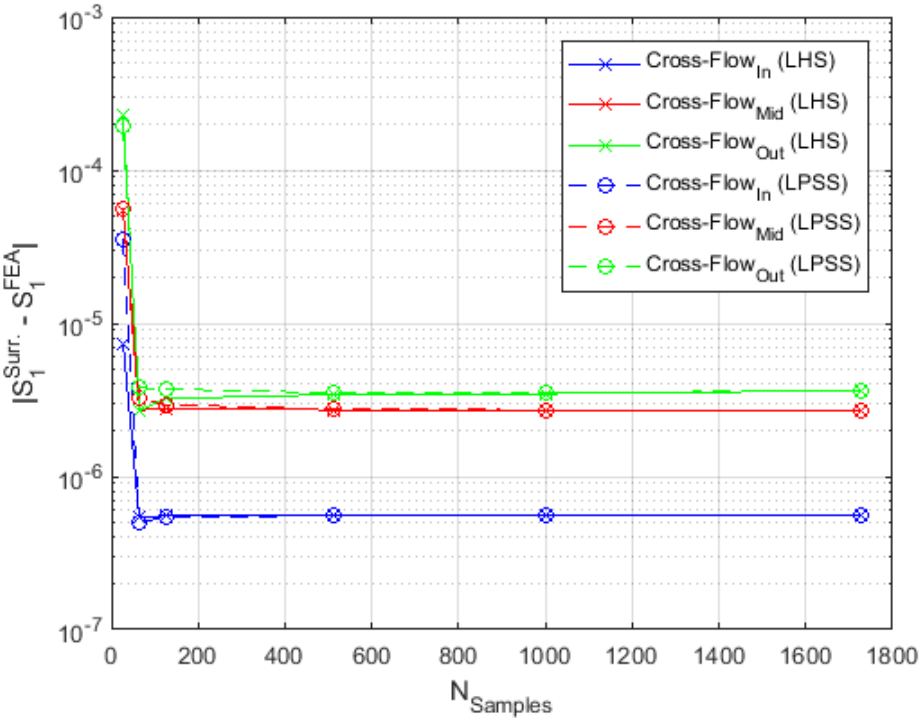

Figure 3.10 Comparison of LHS-Trained and LPSS-Trained Surrogate-based Sensitivities to Benchmark for Cross-Flow Loads

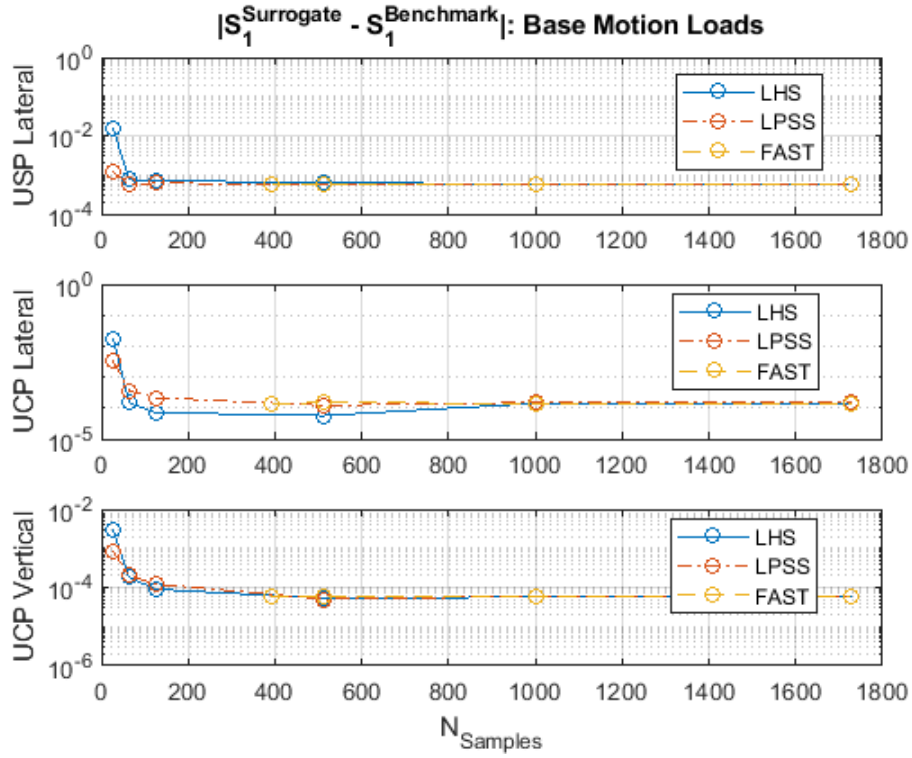

Figure 3.11 Surrogate-based Sensitivities vs. Benchmark for Base Motion Loads

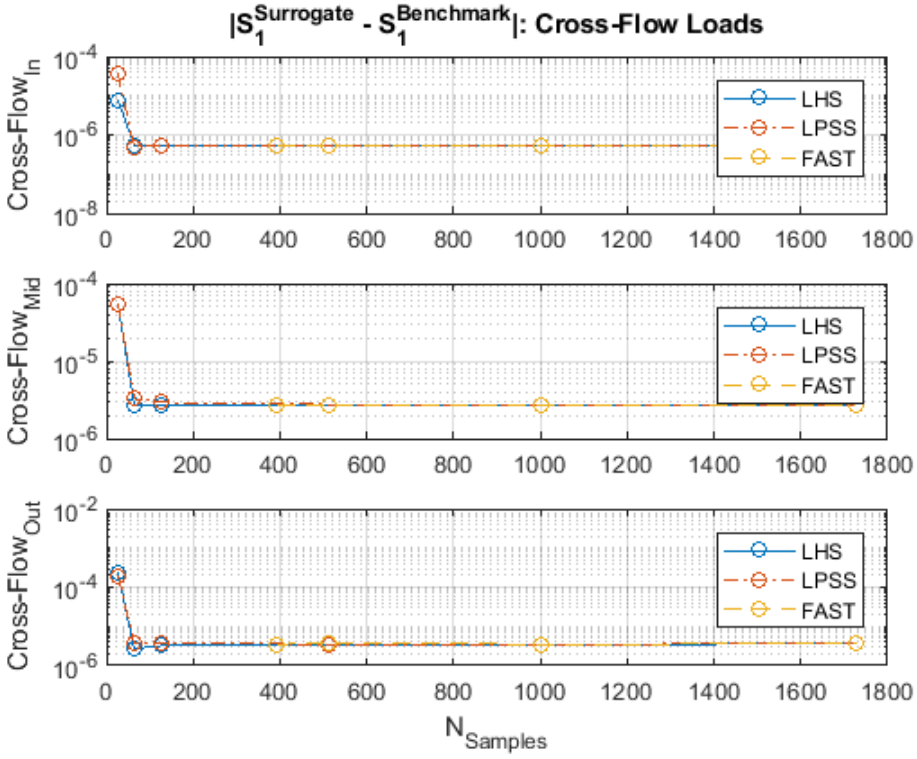

Figure 3.12 Surrogate-based Sensitivities vs. Benchmark for Cross-Flow Loads

3.3 Discussion from Evaluation of UI Flow-Induced Vibration

3.3.1 Surrogate Model Accuracy

It is of interest to determine how the relative difference in sensitivity indices compares with an independent measure of surrogate model accuracy, such as the error between output strains predicted by the Surrogate model and an independent (i.e., not used to train the surrogate models) set of data from the full-order model. To examine this, 100 full-order simulations were randomly sampled from a Uniform distribution independent of the training sets of data used to establish the Surrogate models. Then, the trained Surrogate models were exercised on this independent dataset of 100 to evaluate how well the surrogate model estimations agreed with the full-order model results. The measure of error chosen for these tests was the root mean squared (RMS) error based on the Frobenius Norm, which may be defined as:

$$\|E\|_F \equiv \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left|\left(\epsilon_{FullOrderModel} - \epsilon_{SurrogateModel}\right)_i\right|^2} \tag{54}$$

where $\epsilon_{FullOrderModel}$ and $\epsilon_{SurrogateModel}$ are the i^{th} result predicted by the full-order model and Kriging model, respectively. Variable ϵ represents the axial strain on the upper support skirt. Figure 3.13 shows the RMS error between the full-order model and surrogate models constructed from three different DOE strategies, LHS, LPSS, and FAST sample points, and each with a varying number of training datasets from 27 to 1,728. A decrease in error by approximately a factor of 10 is observed with increasing sample size beyond 64 for all Surrogates and, while some small variability is observed from the Surrogates which were trained from the different DOE methods, the error is less than 0.01 $\mu\epsilon$, and thus judged negligibly small,

in every case for samples sizes of 125 and greater. Given the negligible changes in GSA results above 125 samples, for this stationary FIV problem, there is minimal accuFigure 3.13Figure 3.9 and Figure 3.10 For further comparison, these results were also compared with a Box-Behnken design (Montgomery, 2013) in the interest of understanding how such a traditional DOE approach may perform, and the RMS error of the Box-Behnken design was 15% higher than that associated with LHS, LPSS, and FAST.

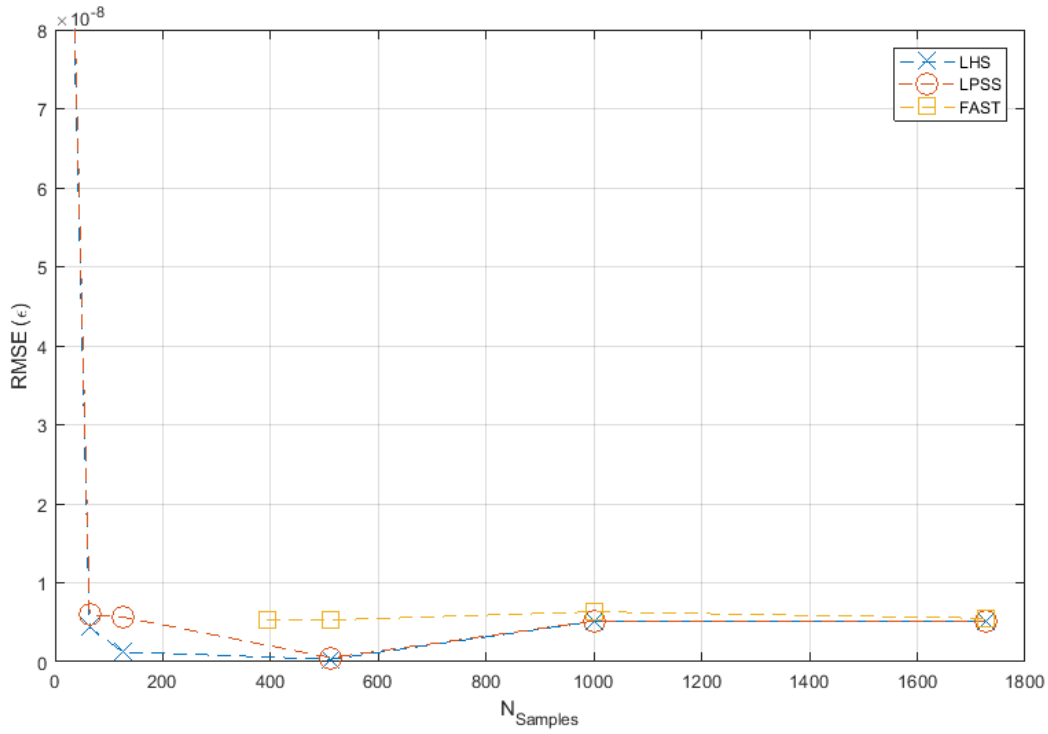

Figure 3.13 Surrogate Model Verification by Comparison to Test Set using Root Mean Squared Error (Strain)

3.3.2 GSA Comparison from Full-Order Model versus Surrogate Models

Although at sample sizes of greater than 500, the differences in surrogate-based GSA results with respect to full-order model GSA results were shown to be small, and also insensitive to the method used to sample the parameter space (e.g., FAST, LHS, or LPSS), some noteworthy differences were observed for smaller sample sizes. First, LHS and LPSS reduced the number of samples needed to assess sensitivity as compared to sampling the parameter space directly with FAST. The minimum number of samples used for FAST was 393, based on the number of integration points needed to resolve the underlying periodic functions for the dimensionality of this problem. In terms of both the surrogate verification error and global sensitivities, LHS and LPSS provide comparable results at 64 samples. Second, for the very small sample size of 27 the error of the global sensitivities determined from the surrogate trained with LPSS is substantially lower than that trained with LHS. As such, although interaction effects may not be strong, the variance reduction provided by LPSS as compared with LHS for a very small number of samples is insightful.

The similar behavior amongst the three methods of sampling the parameter space, as well as the agreement between the surrogate-based sensitivities and those computed from the full-order model, provides confidence in the stability of the results. Furthermore, the agreement between sensitivity indices calculated by the surrogate and full-order models lends credence as to the veracity of the surrogate models. Although the surrogate verification error was quantified in terms of a strain value and the global sensitivity values are unitless, a comparison between these errors may be observed by comparing Figure 3.13 to Figure 3.6 for FAST, and to Figure 3.9 and Figure 3.10 for LHS and LPSS.

One particularly significant physical insight from the results relates to the downcomer. The downcomer forcing function has historically received significant interest (e.g., (Au-Yang & Jordan, 1980)), and to some extent this study helps to substantiate the significance of that particular forcing function. The downcomer forcing function acts directly upon lower internals components, such as the core barrel, and that motion is then coupled with UI components through the upper support flange. From this study, it is apparent that the downcomer forcing function not only directly affects the lower internals response (e.g., core barrel), but also manifests itself as a base motion load on the UI structures and proves even more dominant as a base motion load than the cross-flow loads acting directly upon the UI components. This is a meaningful observation in the sense that much effort has traditionally been devoted to characterizing flow fields in the upper plenum region of the reactor. It is thus apparent that from the perspective of the structural dynamic response of the upper support assembly subjected to flow-induced excitation, rigorous characterization of the upper plenum flow field may be, to some extent, unwarranted. As a caution, this point is not necessarily generalizable to all PWRs, but is nonetheless a meaningful observation for the particular analysis presented herein.

These observations have implications in terms of both nuclear component design and diagnostics. For design of a complex reactor assembly, characterization of forcing functions incurs significant engineering cost involving scale model test programs and computationally expensive computational fluid dynamics (CFD) simulations. Thus, if the forced response of a complex assembly is governed by a select few forcing functions, albeit of uncertain magnitude, a surrogate model defined by those forcing functions may be exercised easily to make risk informed decisions to focus on development of specific forcing functions during the design process. For diagnostics, dynamic instrumentation of an operating reactor incurs great cost to

plant owners. Thus, if a given component is experiencing anomalous behavior, which is observable from its structural dynamic behavior, and it is known which few forcing functions govern the associated forced response, the amount of required dynamic instrumentation may be limited in order to diagnose potential problems. Furthermore, an inverse problem may be constructed and parameterized so as to seek optimal, or most-likely, values of the dominant forcing functions.

3.3.3 Conclusions from Upper Internals FIV Evaluation

A surrogate model prototypic of an UI reactor subassembly was constructed. This model included 6 independent forcing functions (with variance) and one output forced response variable (axial strain). The global sensitivity analyses showed that for the response variable of interests, three of the total six forcing functions dominate the response of the structure. Three different workflows were studied in which Kriging surrogates were trained using different methods of sampling the parameter space; namely FAST, LHS, and LPSS. For large sample sizes, all approaches produced accurate global sensitivities, which provides confidence in the model results and suggests a stable sensitivity analysis result. For relatively small sample sizes, LHS and LPSS were shown to yield surrogates with improved accuracy relative to those yielded from FAST sample points. For very small sample sizes, LPSS is shown to yield improved accuracy relative to LHS.

The optimal selection of Kriging trend and correlation functions depends on the application at-hand. Recently, novel model selection criterion and model averaging technique that employs the information-theoretic multimodel inference have been documented (Sundar & Shields, 2018). Similarly, aggregate surrogate modeling methods adaptively trained by a unique

universal predictive distribution have recently been documented in (Ben Salem, et al., 2017) and (Ben Salem & Tomaso, 2018) initially explored in (Banyay, et al., 2018). Chapters 4.0 and 5.0 involve the application of GSA and Kriging methods, as well as aggregate surrogate modeling, for a reactor assembly model with non-linearities (i.e., non-linear springs and dampers) and non-stationary loading (i.e., loss-of-coolant-accident acoustic or seismic loads).

4.0 Application to Non-Stationary Random Vibration of Reactor System

Reactor assembly system finite element models (SFEMs) are ubiquitous throughout the design and analysis of key systems, structures, and components within PWRs, and not only within Westinghouse (Palamara, et al., 2015), but across the commercial nuclear industry per (Ko & Kim, 2012) and (Ko & Kim, 2013), for example. Despite the centrality of SFEMs to the general mechanical and civil design of such high-consequence facilities, uncertainty quantification of such computational models is generally lacking amongst nuclear industry practitioners. Not performing proper uncertainty quantification can lead to overdesigned components and excessive engineering expenses when it comes to the qualification of key equipment. Furthermore, fiscal pressures within the present commercial nuclear industry are driving engineering organizations to challenge the status quo and develop creative ways by which the promise offered by nuclear energy may continue to be delivered in a cost effective manner. As such, it may be argued that a necessary first step towards an optimal framework for doing computational mechanics in a parsimonious manner is to understand the sensitivity of SFEMs. Thankfully, there is significant precedent for uncertainty quantification in computational structural mechanics applications as evidenced by guidance provided by (ASME, 2006) and deployed in studies such as (Hossain, et al., 2015).

This chapter seeks to present the application of surrogate modeling and GSA methods to a stochastic dynamic finite element analysis of a reactor system. This chapter builds upon the work presented in Chapter 3.0 by additionally considering the following complexities:

- Non-linear response behavior with multiple outputs and more input parameters than previous (described in Section 4.1)
- Non-stationary loading (per methodology described in Section 2.1.2)

4.1 Finite Element Analysis Description

Figure 4.1 and Figure 4.2 show a typical PWR and an associated SFEM, respectively.

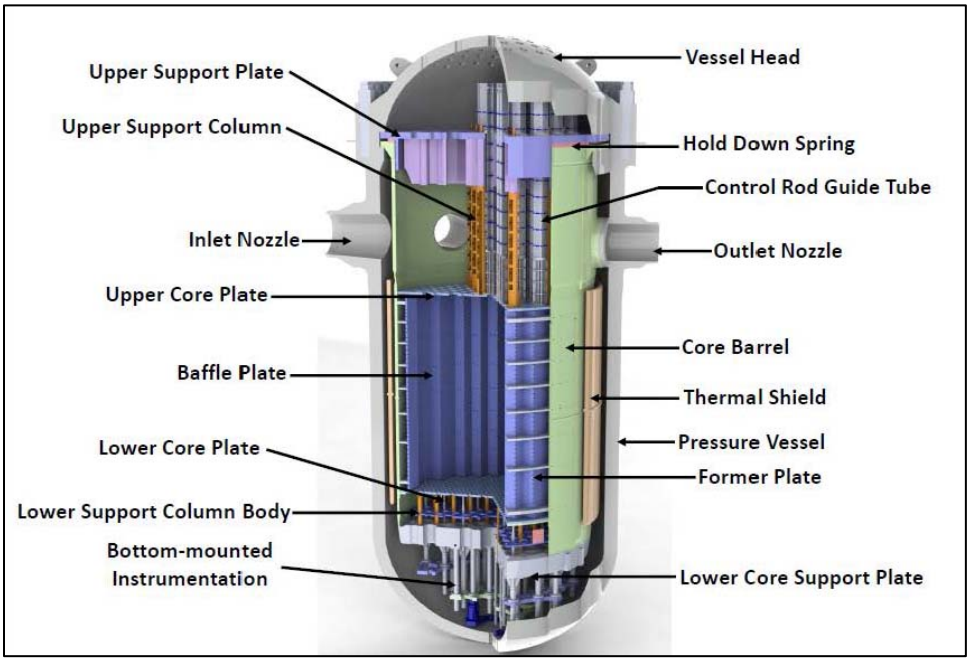

Figure 4.1 Illustration of Typical Pressurized Water Reactor Assembly

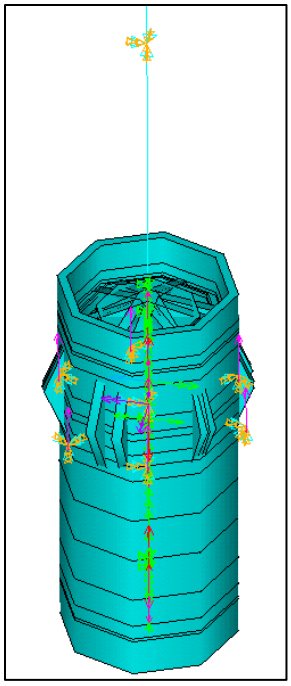

Figure 4.2 Illustration of System Finite Element Model

This model includes approximately 3,000 degrees of freedom consisting of various element types (e.g., beam, shell, spring, damper, etc.). Due primarily to non-linearities associated with the SFEM (Palamara, et al., 2015), such as displacement-dependent spring constants, as well as the non-stationary loading conditions associated with loss of coolant accident (LOCA) and seismic events (required by nuclear design analysis), the SFEM is executed as a transient finite element analysis using implicit time integration as briefly described in Section 2.1.2.

For purposes of the sensitivity analysis, the model was parameterized as follows. The input parameters chosen pertained to model features such as stiffness, damping, gap dimensions, and masses, which vary between various otherwise-similar PWR Reactor Vessel Internals (RVI) structures. In some cases, this variation represents a true change in the underlying physical

parameter, thus constituting an aleatory uncertainty. In other cases, the variation represents the differences imposed to the finite element model in order to bound the plausible scatter, so as to achieve conservative design margins; these basically represent epistemic uncertainties. The output of interest focused upon by this work was lateral acceleration of the Reactor Vessel Closure Head (RVCH).

A preliminary GSA was performed with 16 potentially-important input parameters, and those 16 parameters were down-selected to the six parameters which best account for the variance in the model results of interest. These input parameters are simply identified as P1 through P16 herein, and all pertain to some structural feature of the SFEM such as mass, stiffness, damping, or a gap dimension. The choice to use 6 rather than 16 parameters was intended to more clearly illustrate the surrogate modeling and GSA process for research purposes, although it is recognized that the surrogates could have been trained and this GSA performed using all 16 parameters. The pertinent plots of the coefficient of determination (R^2) and rank correlation coefficient (per Section 2.2.2) for this down-selection are shown in Figure 4.3, and the parameters are briefly described in Table 4.1. Specifically, the output of interest to this analysis is P17 (acceleration of RVCH) and so P4, P10, P11, P15, and P16 constitute the input parameters most strongly correlated with P17. For the GSA runs performed on SFEM, these parameters were varied ±5% about their nominal value (sampled from a Uniform distribution), to account for the typical range of variation on these particular parameters between many operating PWRs.

Table 4.1 SFEM Parameters

Parameter Number	Description [1]	Identifier
P4	Rotational Stiffness at Location 3	$k(\theta)_3$
P10	Damping	ζ
P11	Vertical Gap at Location 5	$g(y)_5$
P12	Number of Components at Location 6	N_6
P15	Stiffness at Location 9	k_9
P16	Stiffness at Location 10	k_{10}

Note:
1. The location identifiers are left intentionally generic in this dissertation so as to protect information which may be considered proprietary. It is judged adequate for purposes of this research to merely identify locations with generic numbers.

Given the six parameters which most influence the model result of interest, a full-factorial DOE was constructed in order to identify statistically significant 2-term interaction effects present amongst those parameters via analysis of variance (ANOVA), which is summarized in Figure 4.4 using the terms described in Table 4.2.

Table 4.2 Analysis of Variance Variability Table Results

Variability Table Result	Explanation[1]
Effect	Describes the size and direction of the relationship between a term and the response variable.
Coef	Describes the size and direction of the relationship between a term in the model and the response variable.
SE Coef	Estimates the variability between coefficient estimates that one would obtain if taking samples from the same population again and again
T-Value	Measures the ratio between the coefficient and its standard error
P-Value	A probability that measures the evidence against the null hypothesis. Lower probabilities provide stronger evidence against the null hypothesis. The null hypothesis is that the coefficient equals 0, which implies that there is no association between the term and the response.

Note:
1. Explanations taken from https://support.minitab.com/en-us/minitab/18/help-and-how-to/modeling-statistics/doe/how-to/factorial/analyze-variability/interpret-the-results/all-statistics-and-graphs/coded-coefficients-table.

Determination of the significance of 2-term interactions is done by first performing ANOVA upon the full-factorial DOE. All possible 2-term interactions between all 6 parameters are included in the ANOVA, and the top four interactions for which the p-value was smallest are identified and highlighted red (indicating their unimportance).

Next, ANOVA was again performed on those results from the full-factorial DOE excluding the terms for which the p-value was greater than 0.1 in order to more clearly see which 2-term interactions carried the greatest statistical significance. For example, the P4 and P16 interaction had a p-value of 0.100 from the first ANOVA performed on the full set of 2-term interactions. When the ANOVA was performed a second time excluding those 2-term interactions with larger p-values, the p-value for the P4 and P16 interaction decreased to 0.087 thus providing some confirmation that 2-term interaction had statistical significance.

Finally, it may be seen that P4 and P16, and P15 and P16 interact, as the two-term p-values are less than 0.10 upon truncating the non-influential terms. The p-value evaluates how well the sample data (i.e., finite element analysis solutions) support the argument that the null hypothesis is true; a low p-value suggests that the sample data provides enough evidence that the null hypothesis can be rejected and is thus a parameter of importance. Note that the p-value of 0.10 was somewhat arbitrarily chosen as a reasonable means of filtering which interactions are influential versus those which are not.

Given the observations of significant 2-term interactions, the subdomains for LPSS were set up as shown in Table 4.3 in a manner analogous to the LPSS design used in Chapter 3.0 (see Table 3.2). That is, as described in Section 2.4.2, variance reduction can be expected from an LPSS design if terms of which interaction effects may be present are grouped together within a given sub-domain. As such, P4, P15, and P16 are all grouped together in one subdomain, and the remaining three parameters a grouped together on a second subdomain.

Table 4.3 Subdomains for LPSS of SFEM

Model parameters	**sub-domain dimension (N_i)**	**strata**	**number of samples**
P4, $k(\theta)_3$P15, k_9 P16, k_{10}	$3^{(1)}$	3 4 5 8 10 12	27 64 125 512 1000 1728
P10, ζ P11, $g(y)_5$ P12, N_6	3		
number of dimensions	6		

Note:
1. Indicated sub-domain established based on observed 2-way interactions between P4 – P16 and P15 – P16 terms from ANOVA.

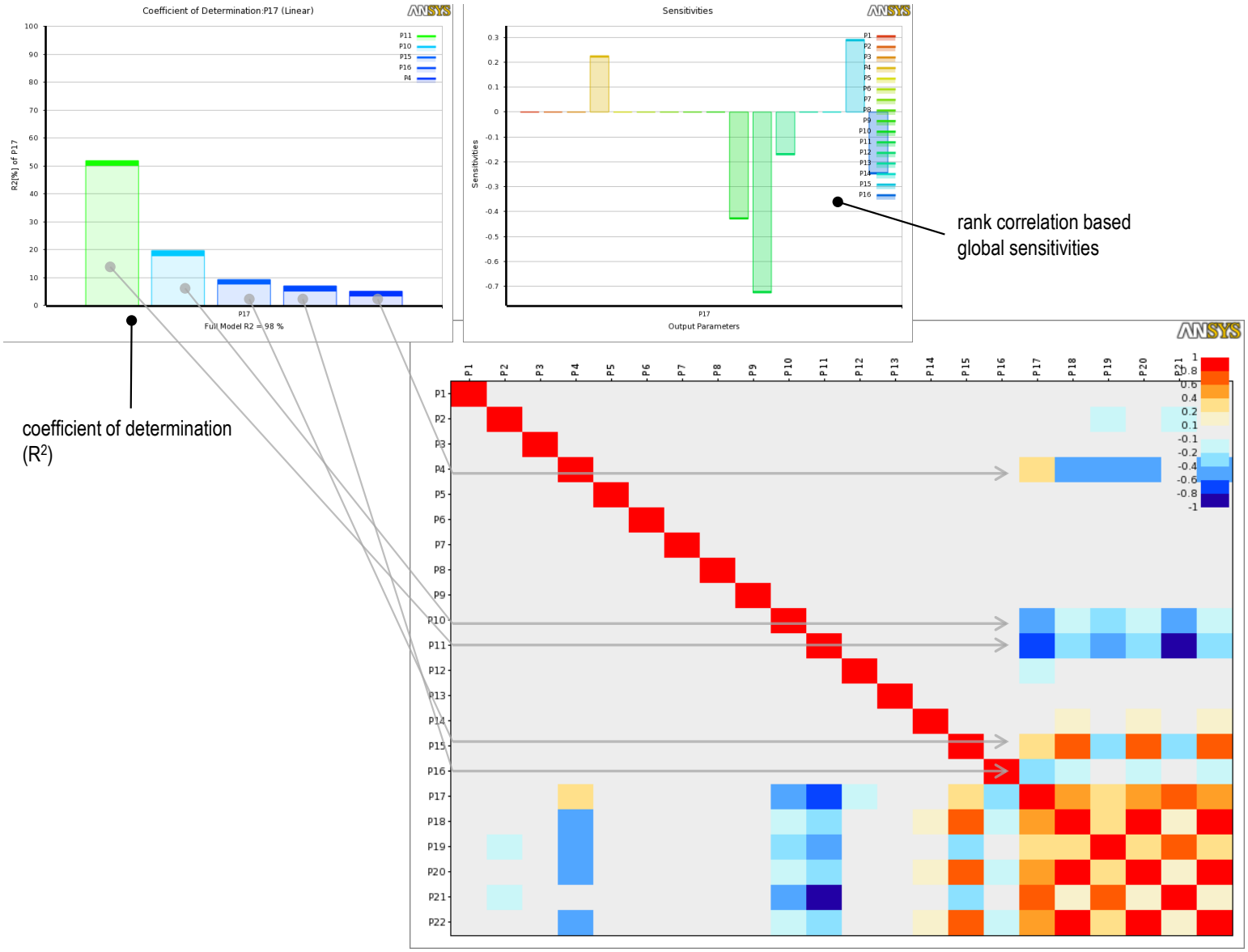

Figure 4.3 Down-Selection of Parameters for SFEM Sensitivity Analysis

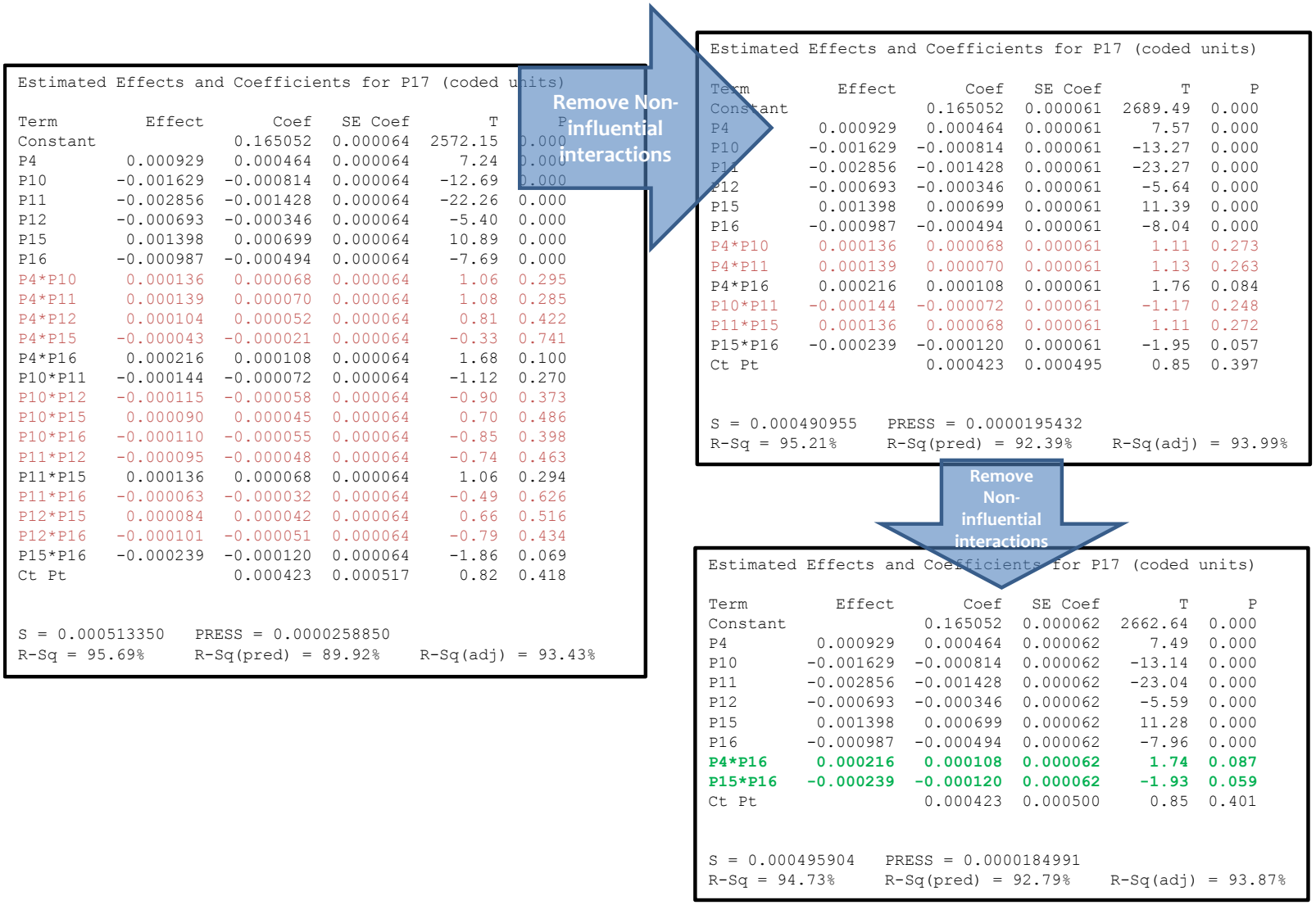

Figure 4.4 Identification of Input Parameter Interactions for SFEM

4.2 Sensitivity Analysis of Non-Stationary Random Vibration for SFEM

Having demonstrated in Chapter 3.0 the successful use of a surrogate model to perform GSA, a similar approach is employed for the SFEM. As before, Kriging surrogates were developed through LHS and LPSS. Sections 4.2.1 and 4.2.2 show the behavior of the SFEM using equivalent methods to those used in Sections 3.2.1 and 3.2.2 for the UI FIV model.

Of preliminary interest is quantification of the percent change in model outputs when a single input parameter is varied. Figure 4-5 through Figure 4-7 show how six model outputs (including RVCH lateral acceleration (P17) and five other acceleration response at different locations/directions) change, relative to their nominal value, with respect to variation of each of three separate input parameters, P6, P8, and P15. It may be seen that the structural dynamic system behavior is indeed non-linear with respect to the range of variation on the three selected input parameters. Furthermore, it may be seen that some model outputs are more sensitive to a given input change, and that the sensitivity of a given model output is significantly different for each input parameter. For example, it may be seen that the percent change for most model outputs shown in Figure 4-7, which pertain to ±10% variation in input P15 significantly exceed the percent change for the same model outputs shown in Figure 4-5 and Figure 4-6, which pertain to ±50% and ±20% variation in inputs P6 and P8, respectively. This implies that input parameter P15 has far higher relative importance than P6 and P8. While such a conclusion would be true if varying one factor at a time, that inference may not necessarily be true for cases in which multiple factors possess uncertainty, and thus could vary from their nominal value. That is, interaction and combinatorial effects could produce significant non-linearities in model

output trends, which would motivate the use of GSA, which surveys the full parameter space considering all possible combinations of input parameters. Such a GSA is presented in the following sections. While Figure 4-5 through Figure 4-7 offer some perspective on the different potential responses of interest, the GSA which follows in this chapter focuses upon the lateral acceleration of the RVCH, since that constitutes a key SFEM output response which drives the stress within numerous sub-components of the reactor system. The RVCH lateral acceleration is identified as parameter P17 herein.

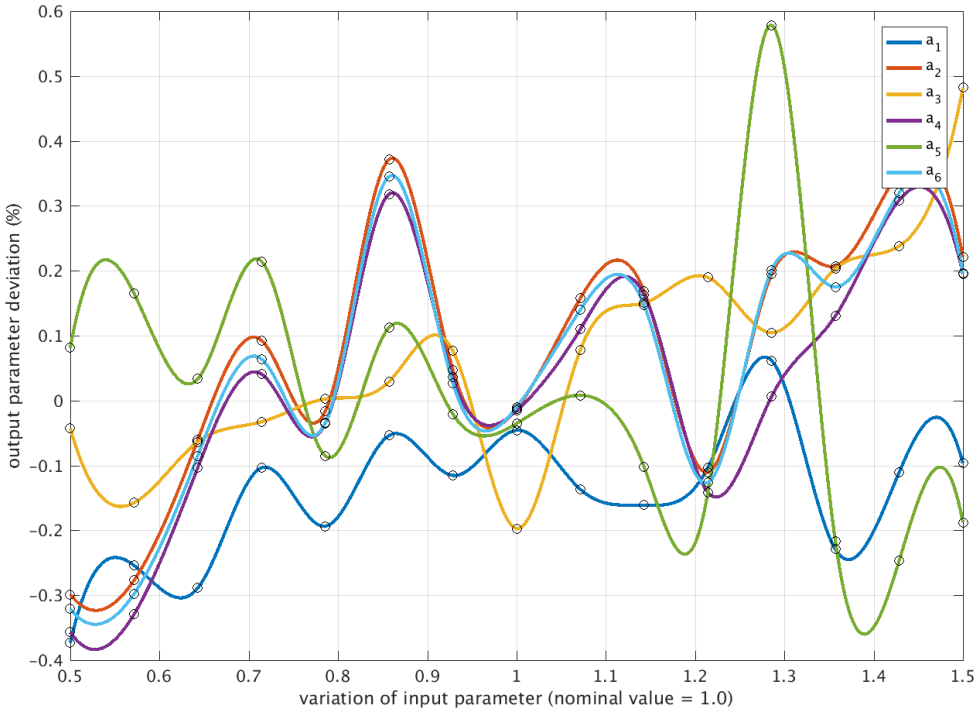

Figure 4.5 Single Factor Sensitivity of 6 Outputs to Input P6

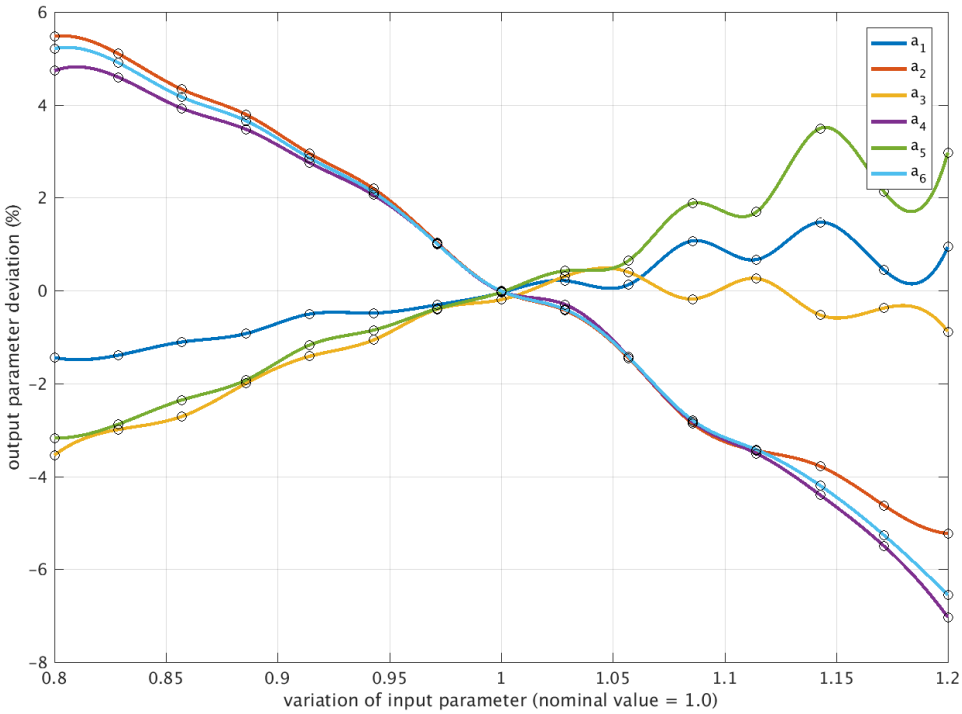

Figure 4.6 Single Factor Sensitivity of 6 Outputs to Input P8

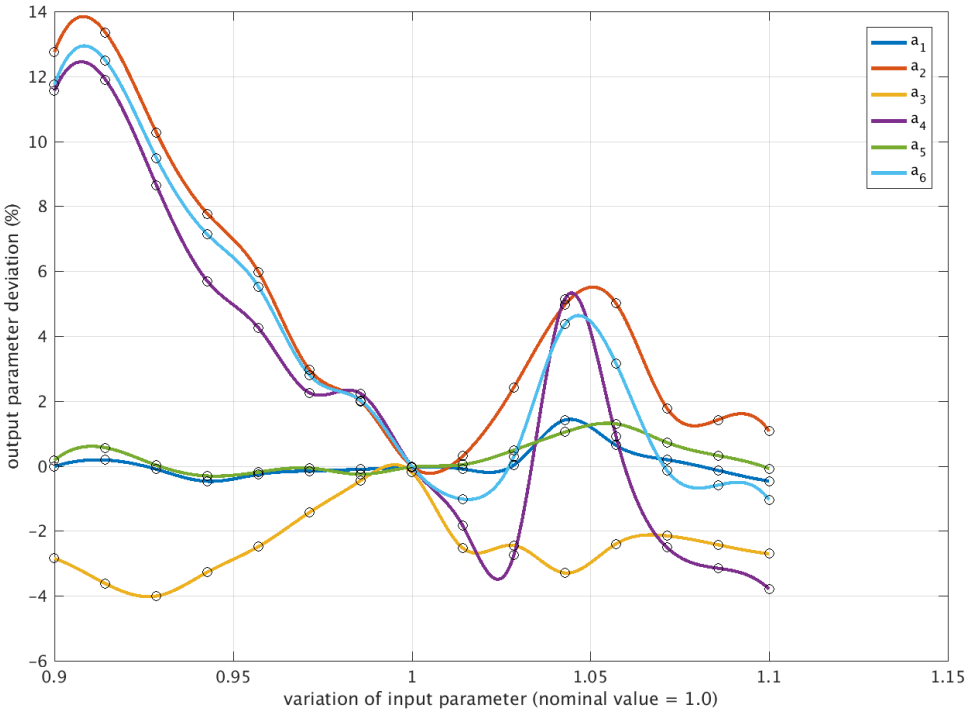

Figure 4.7 Single Factor Sensitivity of 6 Outputs to Input P15

4.2.1 Sensitivity Analysis of Full-Order Model using FAST

Figure 4.8 shows the full-order SFEM GSA results using FAST with 5,000 and 8,000 samples, from which it can be seen that the difference in sensitivities upon changing sample size is negligible. Unlike the GSA results presented in Chapter 3.0 in which three of the six parameters imparted nearly zero variance to the model output of interest, five of the six parameters included in the SFEM GSA contribute at least 5% variance to the model output of interest (P17, lateral acceleration of the RVCH). Referring to Table 4.1 for the parameter identifiers, it may be seen that stiffness between two components in the reactor system K_9 (P15) carries the greatest sensitivity, following by the number of components, N_6 (P12). While damping, ζ, shows a non-negligible sensitivity, it by no means renders the sensitivity of multiple other terms negligible. Terms K_9 and N_6 are local features of the system associated with specific components but impart a significant influence on results of key interest. This insight on the relative sensitivity of damping is important in the sense that damping is often perceived to be a parameter which imparts much greater variance to model results than other terms, which was not the case here.

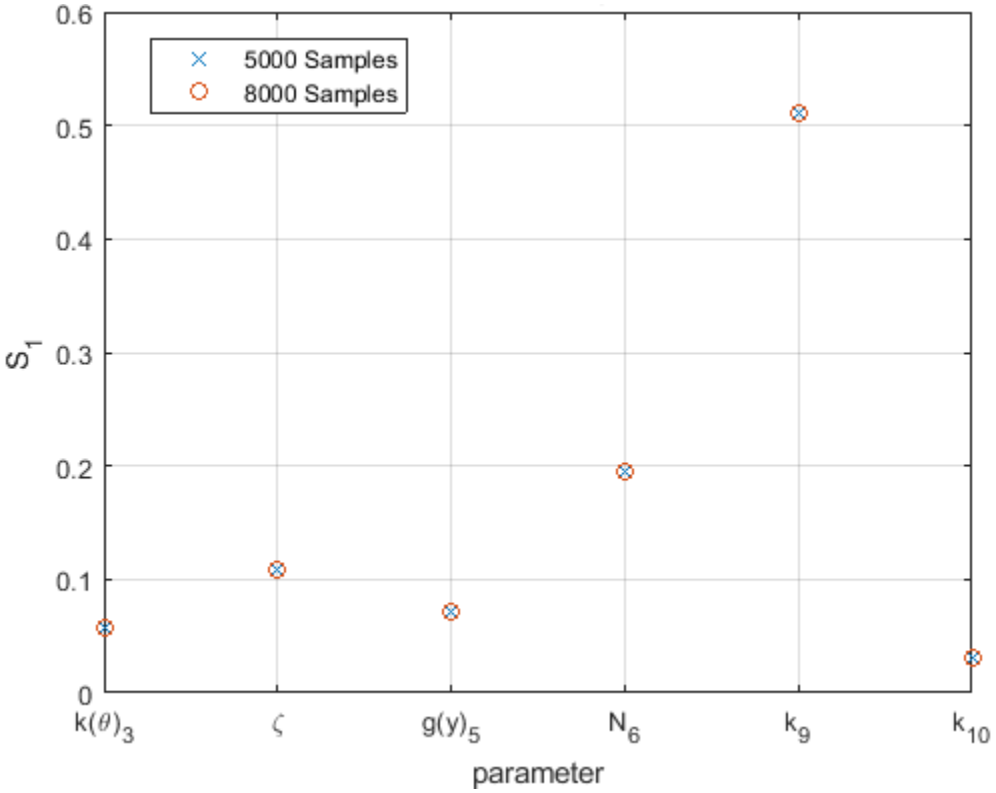

Figure 4.8 Full-Order Model GSA Results for SFEM using FAST

4.2.2 Surrogate-based Sensitivity Analysis

Figure 4.9 shows the Kriging-based GSA results as well as the associated full-order model based GSA results, using both LHS and LPSS. As in Chapter 3.0 , the sample sizes were determined by the stratification of the domains for LPSS, as shown in Table 4.3.

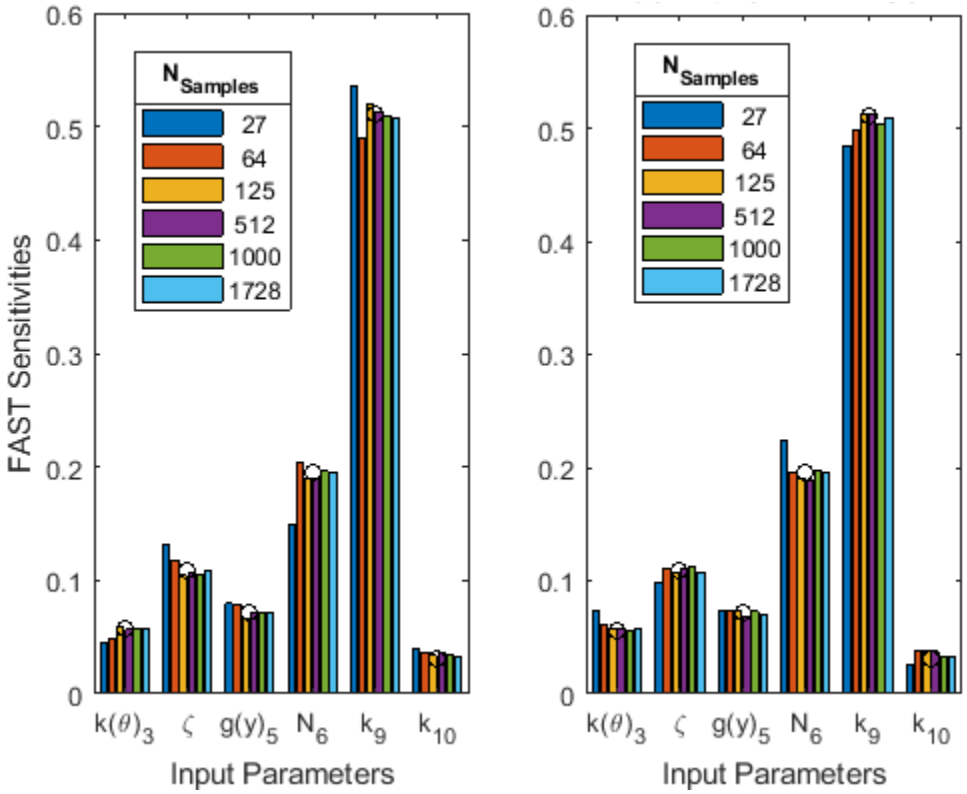

Figure 4.9 Surrogate-based GSA Results for SFEM, Compared to Full-Order GSA Results using LHS and LPSS, Respectively

Figure 4.10 and Figure 4.11 show the difference between the surrogate-based GSA results and those from the full-order model. While the accuracy associated with LHS and LPSS is nearly equivalent at sample sizes of 125 or greater, most LPSS results show notably lower error for sample sizes of 27 and 64. At 27 samples, the parameters for which the GSA error is less for LPSS than LHS include damping ζ (P10), vertical gap at location 5 $g(y)_5$ (P11), number of components at location 6 N_6 (P12), and stiffness at location 10 k_{10} (P16), which was 4 of 6 parameters. This is somewhat interesting in the sense that GSA error was low for the ζ, $g(y)_5$, and N_6 terms which had no notable 2-way interactions from the ANOVA (see Figure 4.4), while greater variance reduction for LPSS may have been expected from terms for which significant

interactions were present. At 64 samples, the parameters for which the GSA error is less for LPSS than LHS include torsional stiffness at location 3 $k(\theta)_3$, damping ζ, vertical gap at location 5 $g(y)_5$, number of components at location 6 N_6, and stiffness at location 9 k_9 (5 of 6 parameters).

Figure 4.12 shows an illustration of the RVCH acceleration response to two of the most influential parameters. The curved shape of this response surface illustrates the value of GSA as opposed to single-factor sensitivities (i.e., the line plots shown in Figure 4-5 through Figure 4-7), as the two terms have a non-linear effect with respect to one another. That is, stiffness at location 9 k_9 (P15) and number of components at location 6 N_6 (P12) have a combined effect on the RVCH acceleration (P17) which changes in a different direction depending on the particular magnitudes of each parameter, and which could easily be misunderstood if all factors were held constant except for one. This interaction effect is further annunciated when considering the combined influences that stiffness at location 9 k_9 (P15) and stiffness at location 10 k_{10} (P16) have on the same result, as shown in Figure 4.13 and Figure 4.14, where the effect that P16 imparts on output P17 depends on the value of P15.

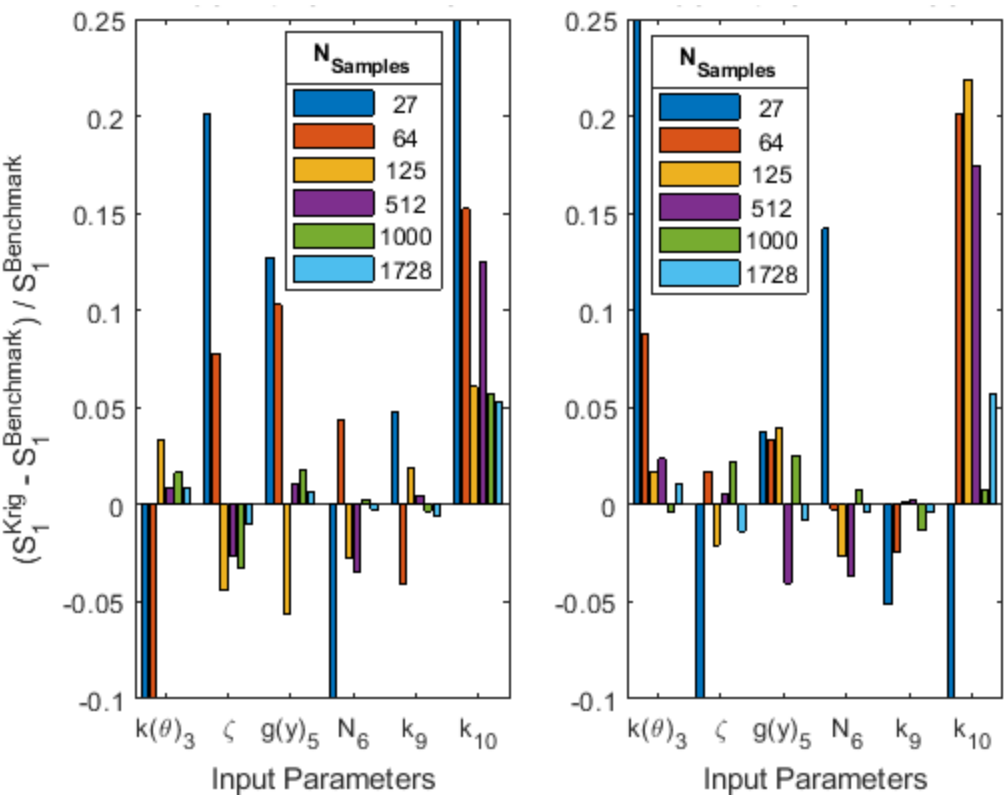

Figure 4.10 Difference between Kriging-based and Full-Order GSA Results of SFEM (Bar Graph) using LHS and LPSS, Respectively

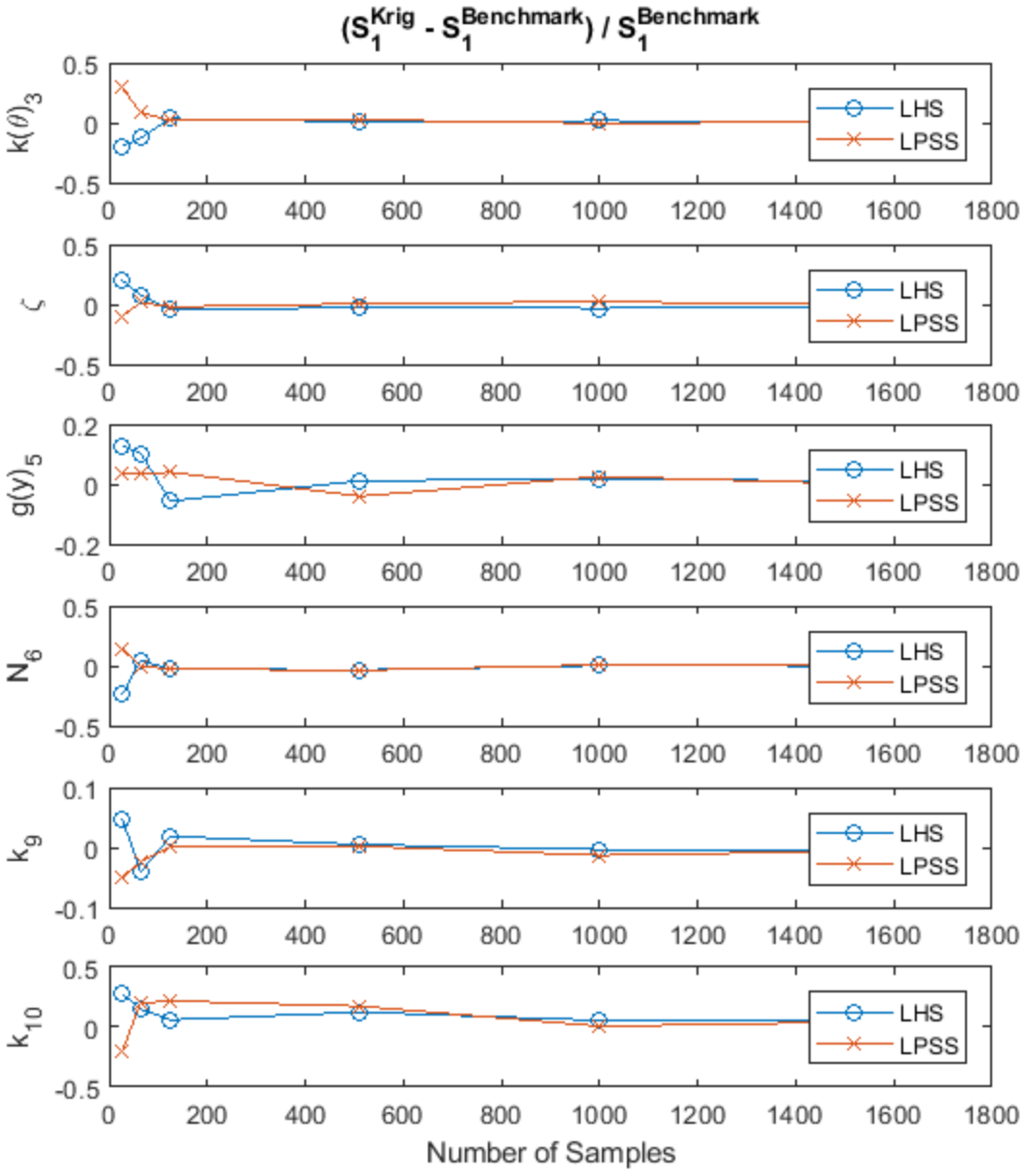

Figure 4.11 Difference between Kriging-based and Full-Order GSA Results of SFEM (Line Graph)

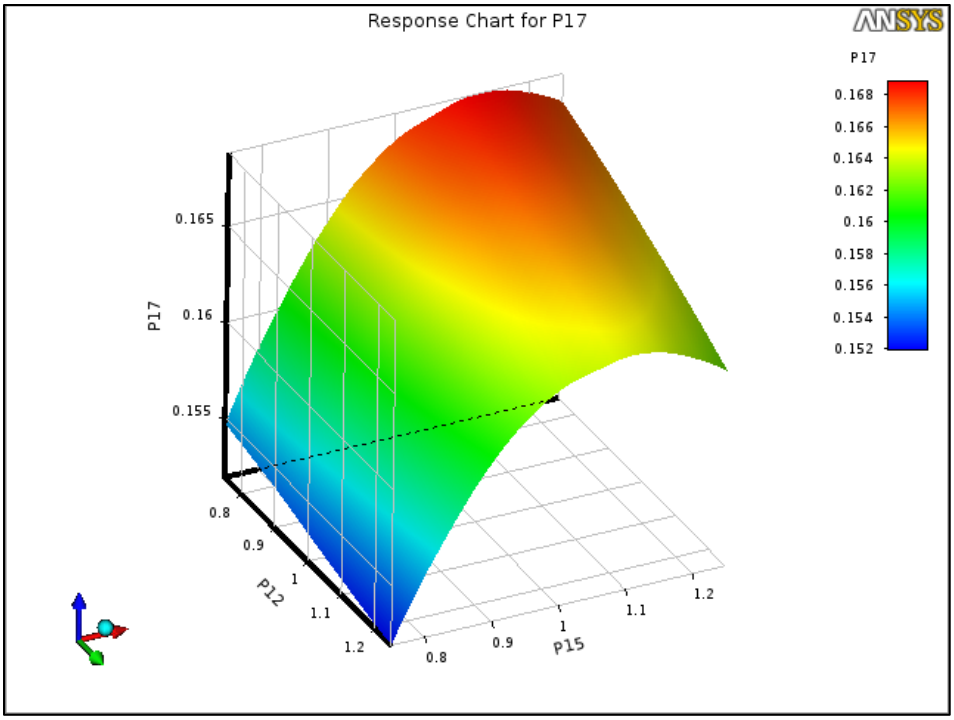

Figure 4.12 Illustration of Kriging Surface for Output P17 with respect to N_6 (P12) and k_9 (P15)

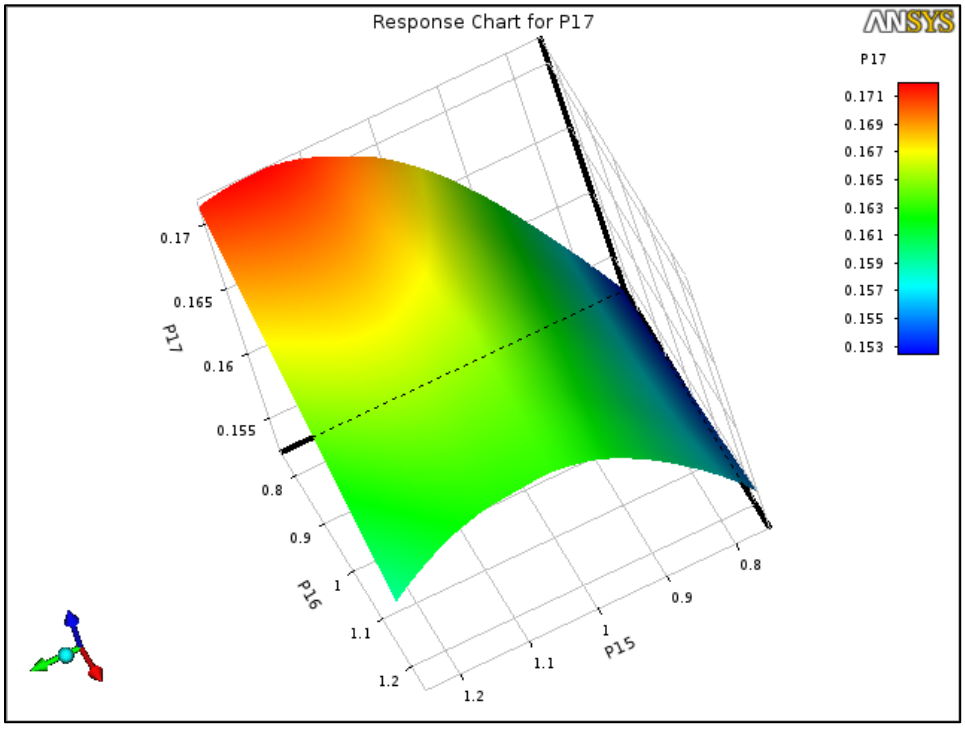

Figure 4.13 Illustration of Kriging Surface for Output P17 with respect to k_{10} (P16) and k_9 (P15)

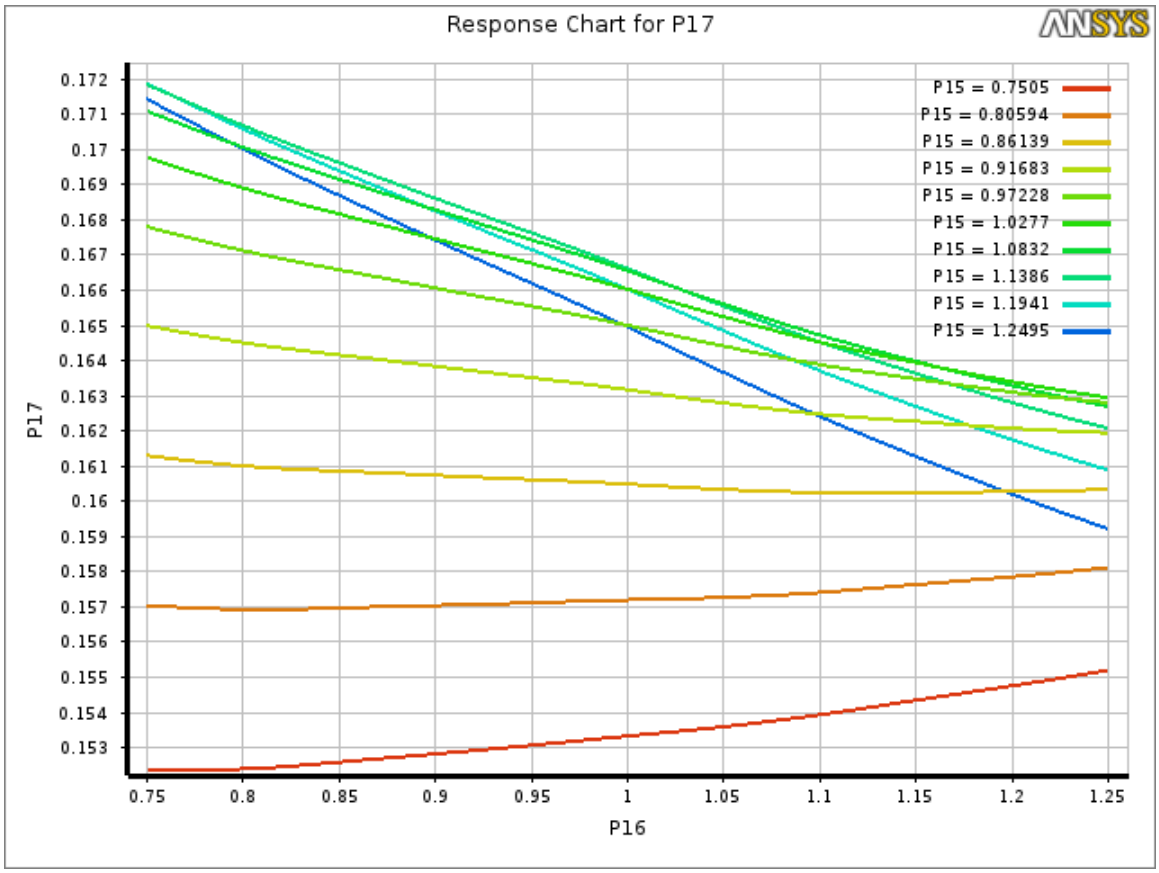

Figure 4.14 Slices through Kriging Surface for Output P17 with respect to k_{10} (P16) and k_9 (P15)

4.2.3 Comparison Studies

To show the changes in sensitivity analysis results due to variations in the amount by which input parameters are varied, as well as the manner in which the dynamic results are post-processed, the next series of studies re-evaluated the GSA with two variations in the SFEM analysis problem structure.

First, although the ±5% range about which the input parameters were varied initially is based upon typical plant-to-plant variations within the operating fleet of PWRs, this is recognized to be a relatively small range. For example, as operating plants age, degradation mechanisms (i.e., wear, embrittlement) manifest themselves in which case some model parameters could depart further from the nominal values (Barker, 2017). Therefore, Section 4.2.3.1 explores the surrogate-based sensitivity analysis for the case in which the input parameter ranges are varied by ±25%.

Second, while the maximum acceleration from the transient analysis is used in practice for ASME Code design analysis of reactor equipment, and is therefore the model output of interest chosen to evaluate within Sections 4.2.1 and 4.2.2, the magnitude of the acceleration response spectrum is also of practical engineering interest. Particularly, for a SFEM which is used to develop loading for detailed sub-model stress evaluations (i.e., reactor components not explicitly resolved in the SFEM itself), spectral amplitudes are important if a particular sub-model is attached to the system and has a resonant frequency coincident with the amplified portion of the acceleration response spectrum. Therefore, Section 4.2.3.2 explores the surrogate-based sensitivity analysis for the case in which the output of interest is the amplitude of the first vibration mode observed from the acceleration response spectrum.

4.2.3.1 Effect of Increasing Input Parameter Variance

The same subset of six parameters was chosen to explore for the sensitivity analyses with the larger ranges of perturbation: rotational stiffness at location 3 (P4), damping (P10), vertical gap at location 5 (P11), number of components at location 6 (P12), and stiffness at location 9 (P15), for which a brief description is provided in Table 4.1. The resultant interactions plot is shown in Figure 4.15, from which it may be seen that interaction effects are much more pronounced for the larger range of perturbation as compared with the companion plot generated from the ±5% case per Figure 4.16, in the sense that the slope of the lines between the two interacting parameters diverge. For example, focusing upon the P4 – P15 interaction, it may be seen that the slope the change in RVCH lateral acceleration P17 when P15 is varied ±25% is larger for when P4 is at 25% above its nominal value (red dashed line) as compared to when P4 is held at 25% below its nominal value (blue solid line).

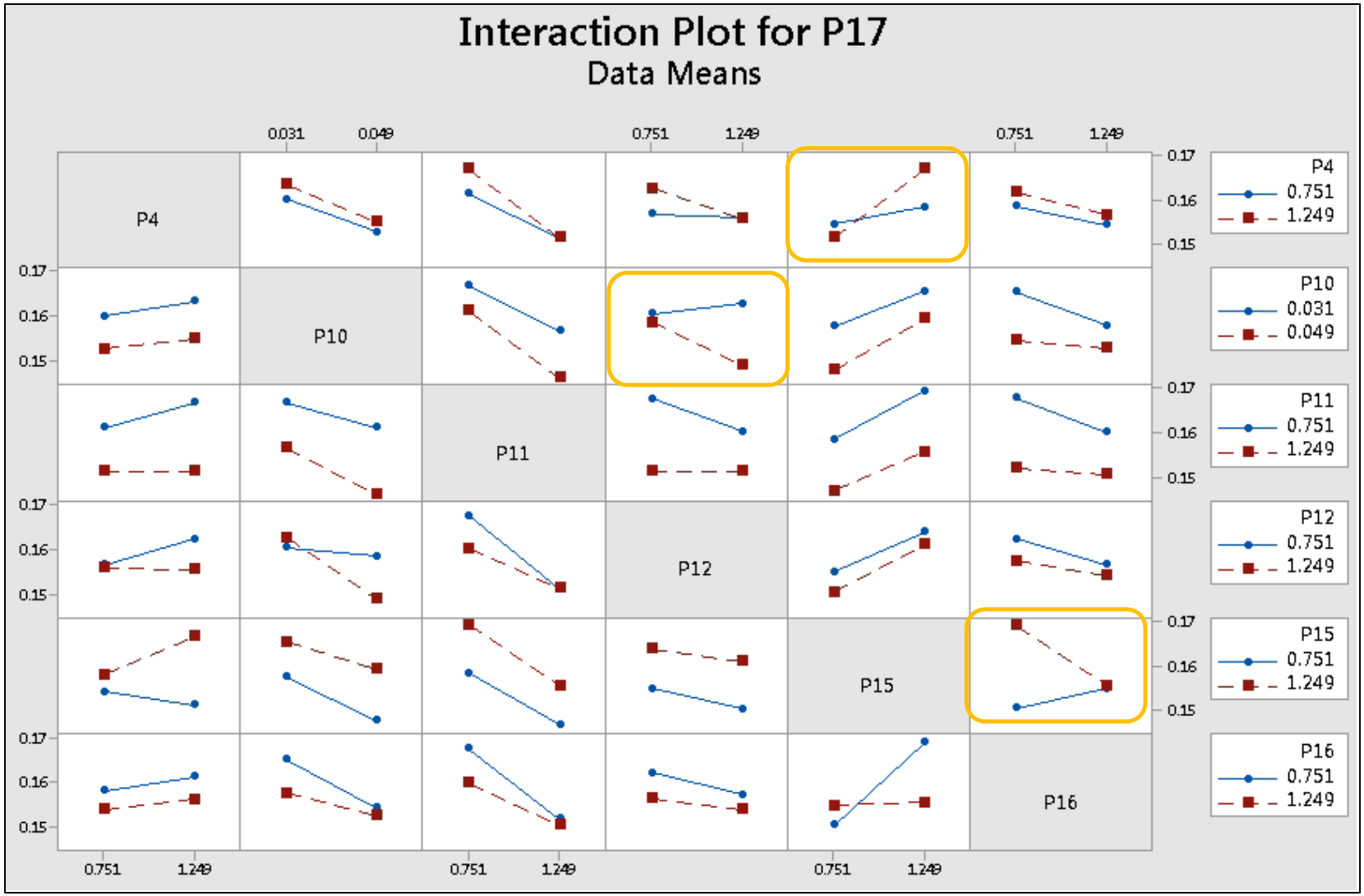

Figure 4.15 Interaction Plot for Lateral Acceleration Response with ±25% Variation

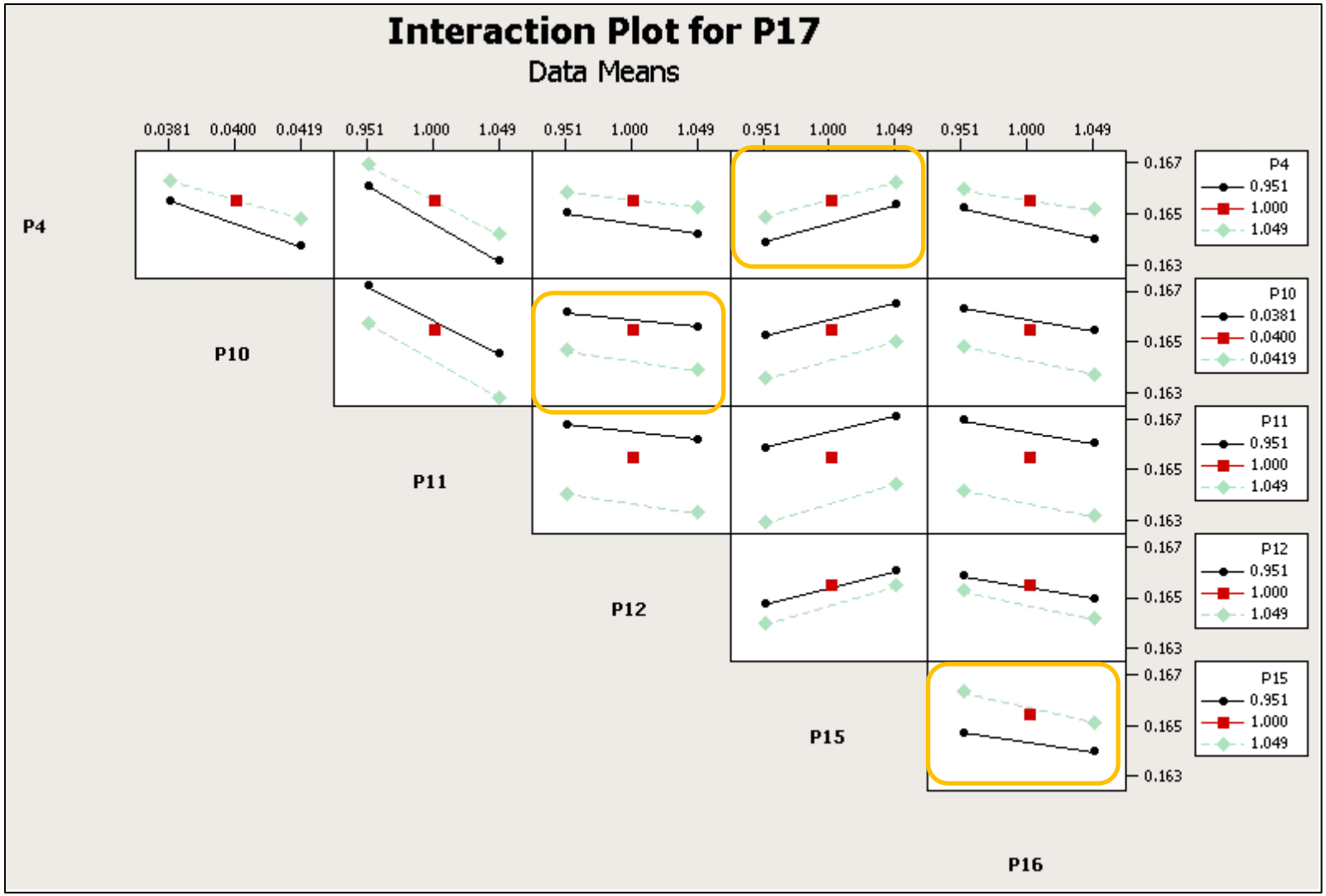

Figure 4.16 Interaction Plot for Lateral Acceleration Response with ±5% Variation

The first-order global sensitivities for the SFEM in which each of the six input parameter were varied by ±25% are shown in Figure 4.17, which is based on drawing 12,000 samples of the full-order model using FAST. These global sensitivities may be compared with those for the SFEM in which each of the six input parameters were varied by ±5% per Figure 4.8. For the analysis with higher variance, parameter k_9 remains the parameter of highest importance, but the relative importance of parameters ζ and N_6 become reversed by increasing the range of perturbation. This suggests that stiffness at a particular location within the SFEM remains the dominant parameter independent of the range of perturbation, but that damping may impart a more significant effect on key SFEM outputs as the range of perturbation is increased. It makes some intuitive sense that as the variation in damping is increased, it could become more influential with respect to other terms. It is also somewhat interesting that the non-influential terms, $k(\theta)_3$, $g(y)_5$, and k_{10}, all were shown to be even less influential for the ±25% variation case in that the first order sensitivities were lower.

Upon inspection of Figure 4.17, it may be seen that the sum of the first-order sensitivities for 8,000 samples falls below 1.0. As the first-order indices represent the main effect contribution of each input factor to the variance of the output, this could suggest the presence of interaction effects, and therefore nonzero higher-order sensitivities, for the situation that the range of parameter variation is increased from ±5% to ±25%. As discussed in Section 4.4 and 4.5 of (Saltelli, et al., 2008), Sobol's method could be used to quantity the interaction effects. Using S to denote a sensitivity index per Section 2.2, the sum of the first-order and higher order sensitivities indices can be expressed as:

$$\sum_i S_i + \sum_i \sum_{j>i} S_{ij} + \sum_i \sum_{j>i} \sum_{l>j} S_{ijl} + S_{123\ldots k} = 1 \tag{55}$$

Although these higher-order indices could be computed (e.g., using UQ Lab (Marelli & Sudret, 2014)), it may be seen from Figure 4.17 that, when using 12,000 samples, the sum of first-order sensitivities does equal 1.0. Therefore, although from 8,000 samples it could appear as though there may be higher-order sensitivities, in fact, the sensitivity analysis had not used a sufficient number of samples to obtain accurate sensitivities.

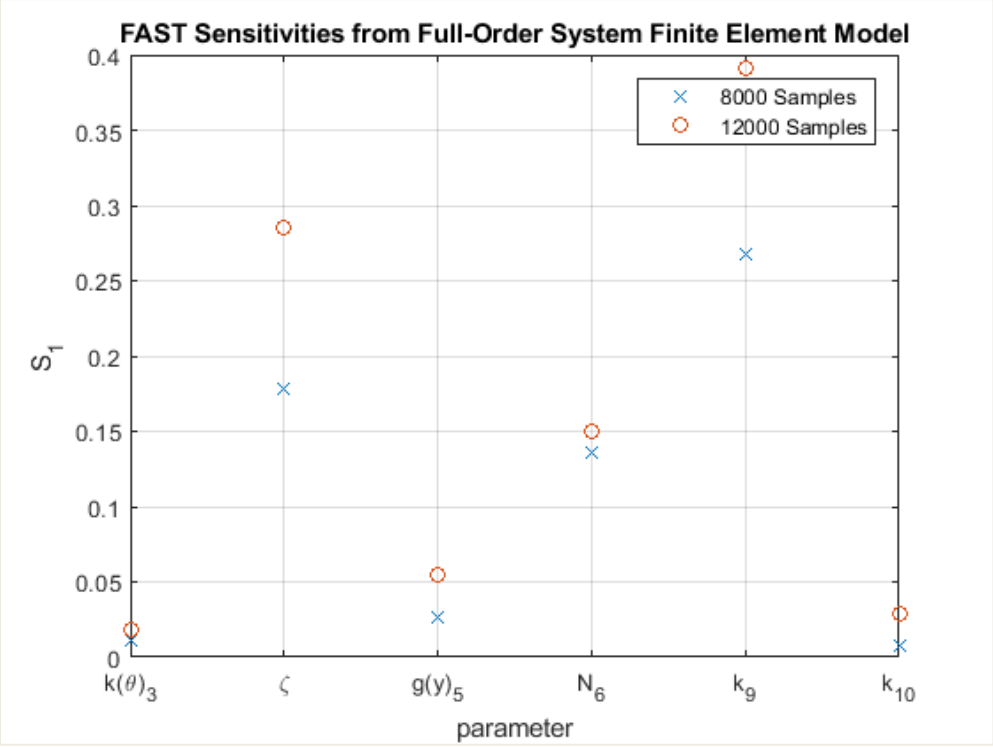

Figure 4.17 Full-Order SFEM Global Sensitivities with ±25% Parameter Variance

Figure 4.18 shows the GSA error with respect to increasing sample size for LHS and LPSS trained Kriging surrogates. At 27 samples, the GSA error associated with LPSS is less than that for LHS for parameters $g(y)_5$, N_6, and k_9 (3 of 6 parameters). At 64 samples, the GSA

error associated with LPSS is less than that for LHS for parameters ζ, $g(y)_5$, N_6, and k_{10} (4 of 6 parameters). Thus, the LPSS error is more consistently lower than the LHS error for the ±5% case (see Figure 4.11) than for the ±25% case (see Figure 4.18). Therefore, the effectiveness of LPSS in developing a surrogate for which the GSA error is smaller than that provided by LHS is slightly less observable when the parameter perturbations are increased. This is somewhat surprising in that greater interaction effects were present for the ±25% case than the ±5% case (see Figure 4.15 compared to Figure 4.16), which might lead one to think that LPSS would impart greater variance reduction compared to LHS and thus result in lower GSA error, but that is not necessarily the case for the SFEM analysis.

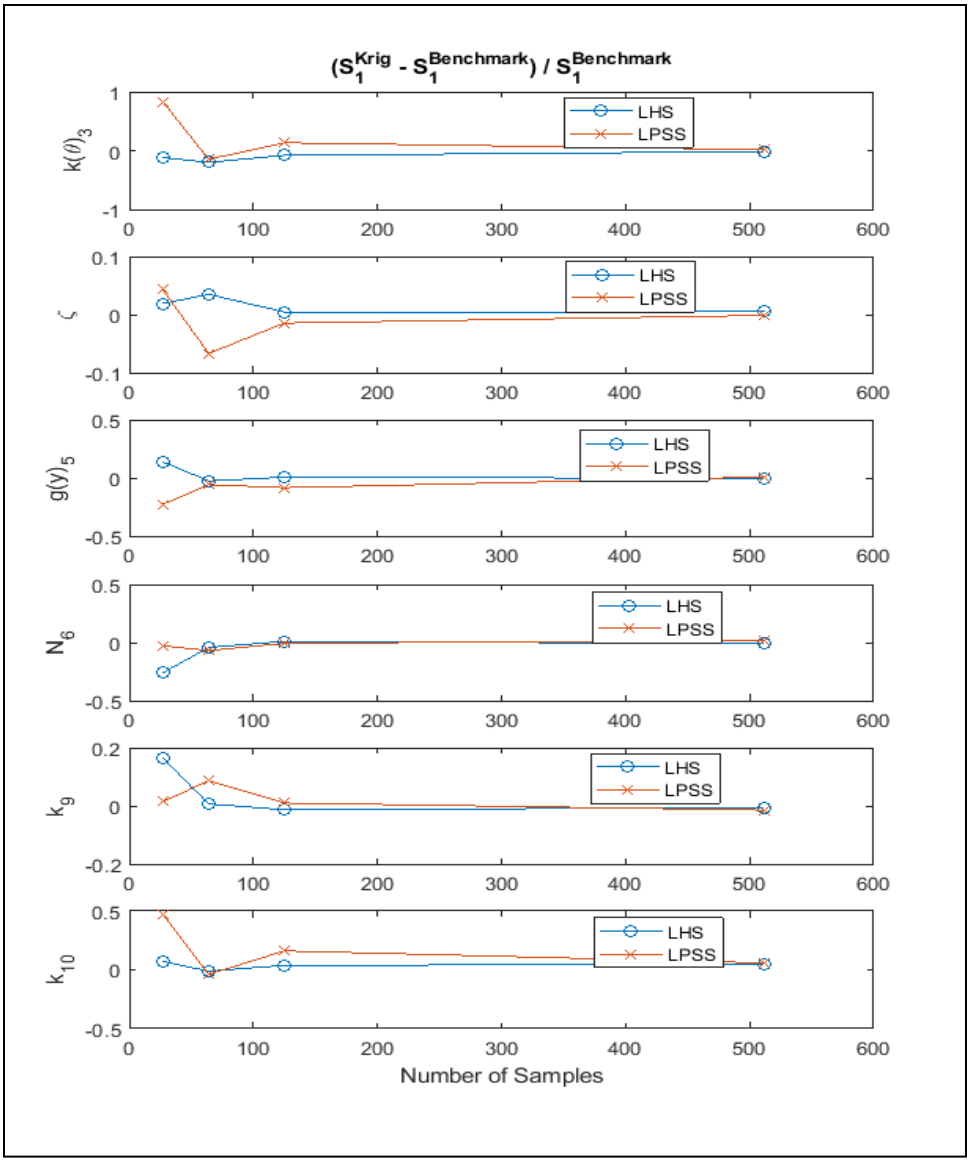

Figure 4.18 Convergence of Surrogate-based Global Sensitivity Indices for System Finite Element Model with ±25% variation on input parameters

In order to better understand these results, a full factorial computational DOE was performed on the SFEM in which the model parameters were varied ±25%. The DOE results are summarized in Table 4.4 within which the parameters with a p-value less than 0.10 are highlighted in the center column pertaining to lateral RVCH acceleration. From this, it may be

seen that only two of the six parameters used from the ±5% variation case show a p-value of less than 0.10, P11 and P12, while the other four parameters had noticeably larger p-values. The factorial DOE p-value differences between the ±5% and ±25% variation cases are consistent with the prior observations of differences between Figure 4.8 and Figure 4.17, in that both measures show that the range of perturbation can change the degree to which variance in the result of interest (e.g., lateral RVCH acceleration) is influenced by changes in particular parameters.

The third (right-most) column of Table 4.4 shows the results from the same full-factorial DOE as was done for lateral RVCH acceleration, but for vertical RVCH acceleration. Here again, the parameters with a p-value of less than 0.10 are highlighted, but it may be seen that those parameters are different from those of low p-value for lateral acceleration. Between the two outputs, of those terms for which the p-value was less than 0.10, only two parameters are shared (vertical gap at location 5 (P11), number of components at location 6 (P12)), while all other parameters with a low p-value were found to be non-influential to the other result. This difference in what was significant suggests that the parameters which could be truncated from the full set due to lack of significance, and the parameters which could be grouped together for LPSS, would change based on the model output of interest.

Table 4.4 Factorial DOE Result for Pressure Vessel Acceleration with ±25% Input Parameter Variation

Input Parameter	P-Value	
	Lateral RVCH Acceleration (P17)	Vertical RVCH Acceleration (P18)
P1	0.000	0.697
P2	0.889	0.000
P3	0.198	0.926
P4	0.263	0.000
P5	0.057	0.359
P6	0.110	0.590
P7	0.118	0.293
P8	0.948	0.119
P9	0.483	0.763
P10	0.385	0.003
P11	0.000	0.063
P12	0.000	0.035
P13	0.014	0.781
P14	0.710	0.013
P15	0.684	0.000
P16	0.000	0.152

4.2.3.2 Evaluation of Acceleration Response Spectra Variation

Figure 4.19 provides a typical acceleration time-history response. The acceleration value of interest from this plot is the maximum value which occurs at approximately a time of 1.07 seconds, and this maximum acceleration over time quantity represents the output of interest studied up to this point of the present Chapter.

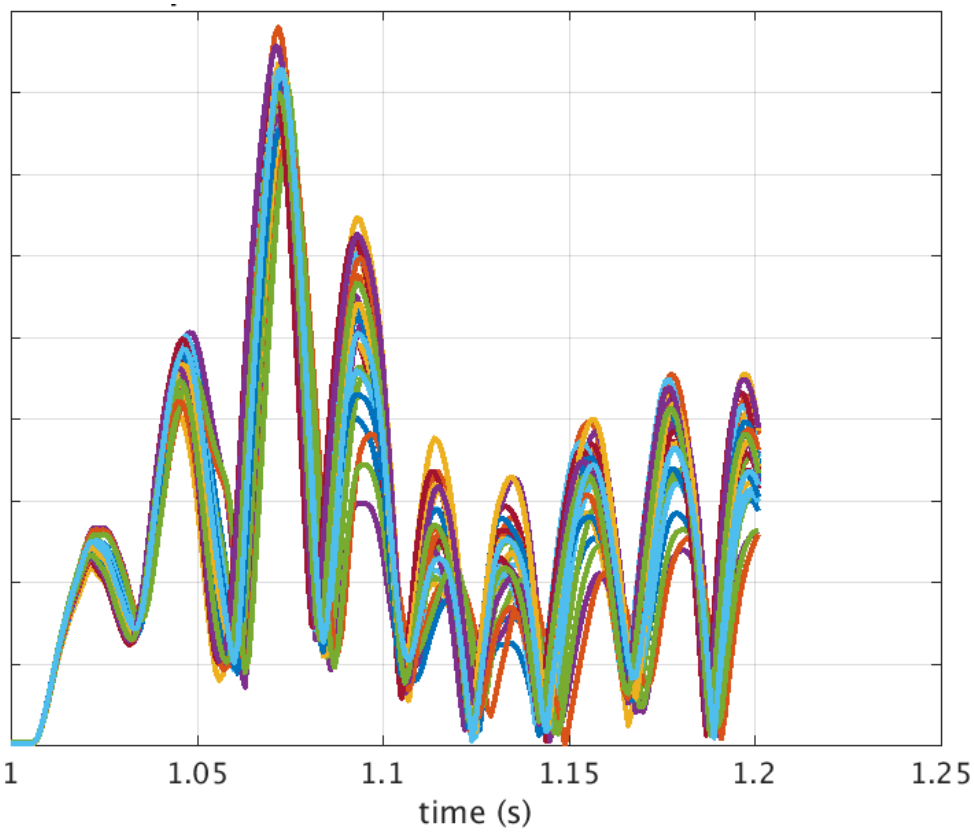

Figure 4.19 Acceleration Time Histories for RVCH lateral motion

Whereas the majority of the SFEM analyses presented in this dissertation focus upon an output which is a maximum over time, this section seeks to explore the performance of these same methods but with respect to an output from the acceleration response spectrum (per Section 2.1.2.2). The maximum acceleration over the transient is based on the superposition of multiple vibration modes at a given moment in time, so focusing on one particular mode of the response spectrum was found to provide different insights on the surrogate-based GSA behavior.

In some situations, the magnitude of the acceleration response at a particular frequency may be of greater interest than the maximum acceleration over the transient. For example, if a component attached to those represented by SFEM has a particular resonant frequency of concern. Examples of such components may include the control rod drive mechanism (CRDM), and piping attached to the inlet or outlet nozzles, as shown in Figure 4.21, as well as components of a simplified head assembly (SHA) as shown in Figure 4.20. Such components are not specifically resolved in the SFEM, but the SFEM is used to generate dynamic loads which constitute forcing functions to subsequent structural evaluations of these components. Therefore, particular resonant frequencies of components associated with CRDMs, SHA, or piping, for example, could respond with significant dynamic amplification when exposed to particular frequencies of the SFEM acceleration response spectrum.

For evaluating the dynamics at a resonant frequency of interest, it is insightful to plot the dynamic result as acceleration response spectra, as shown in Figure 4.22 as an overlay or in Figure 4.23 as a contour. Note that Figure 4.19, Figure 4.22, and Figure 4.23 were generated from the LHS 27-sample run, for illustration purposes. Even for such a small number of samples, the variation on acceleration amplitude in Figure 4.22 is clearly observable and, on a

relative basis, more exaggerated than the variation in the peak acceleration over time per Figure 4.19.

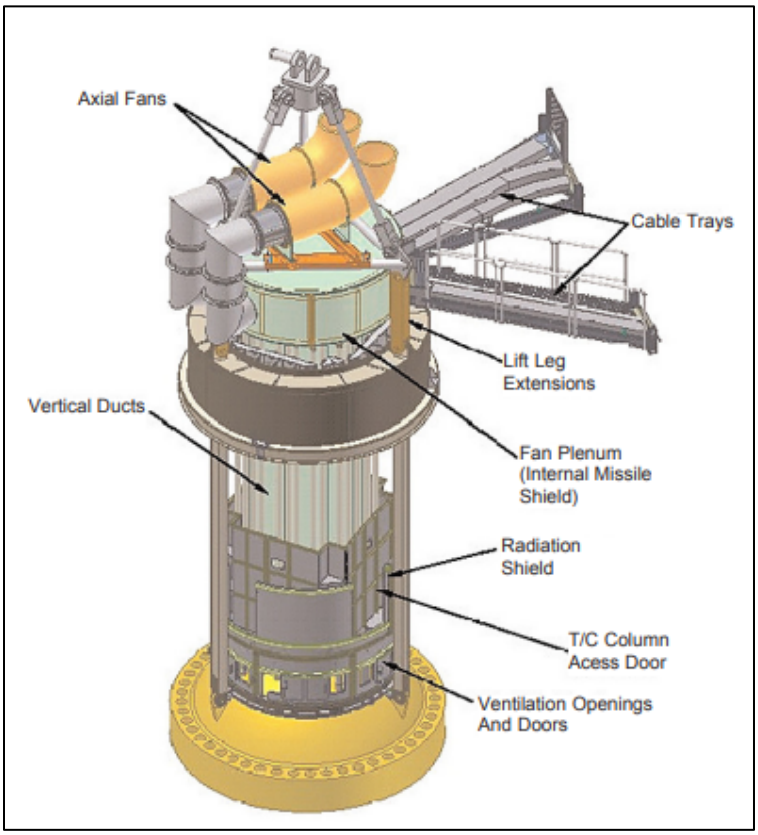

Figure 4.20 Reactor Vessel Closure Head with Simplified Head Assembly

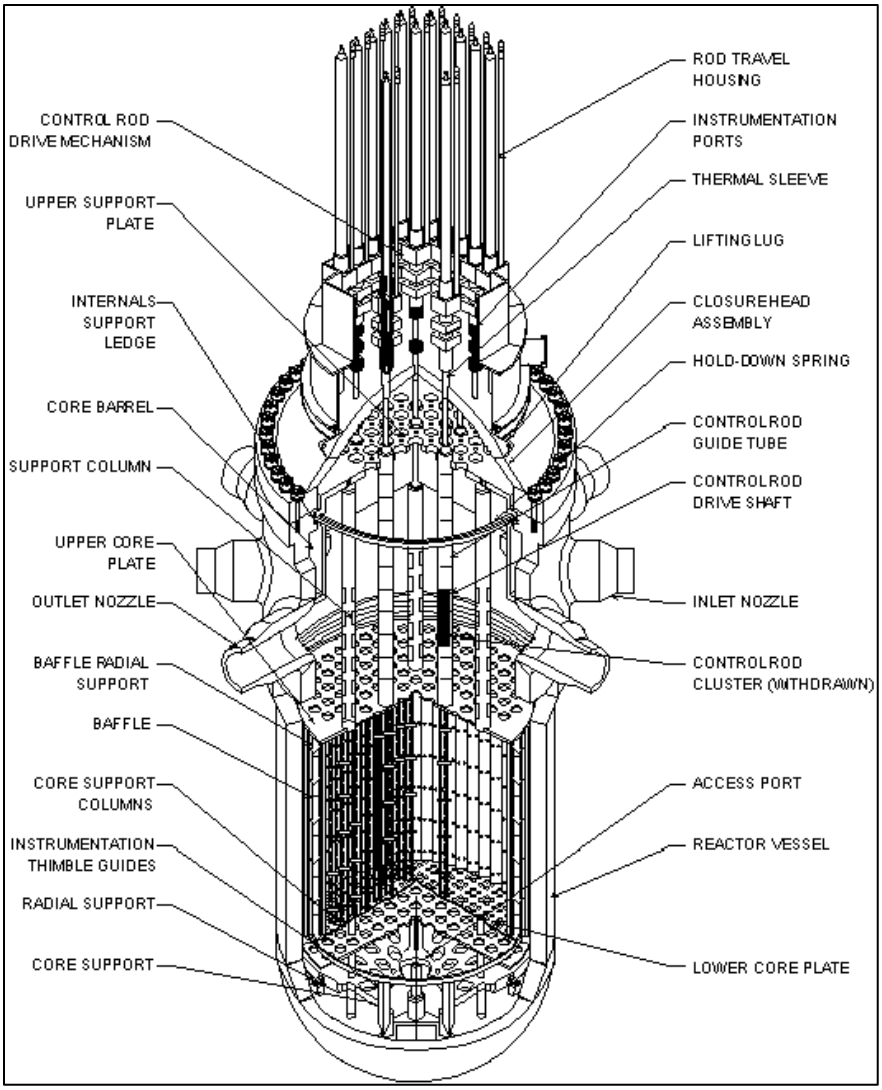

Figure 4.21 Typical Pressurized Water Reactor Schematic

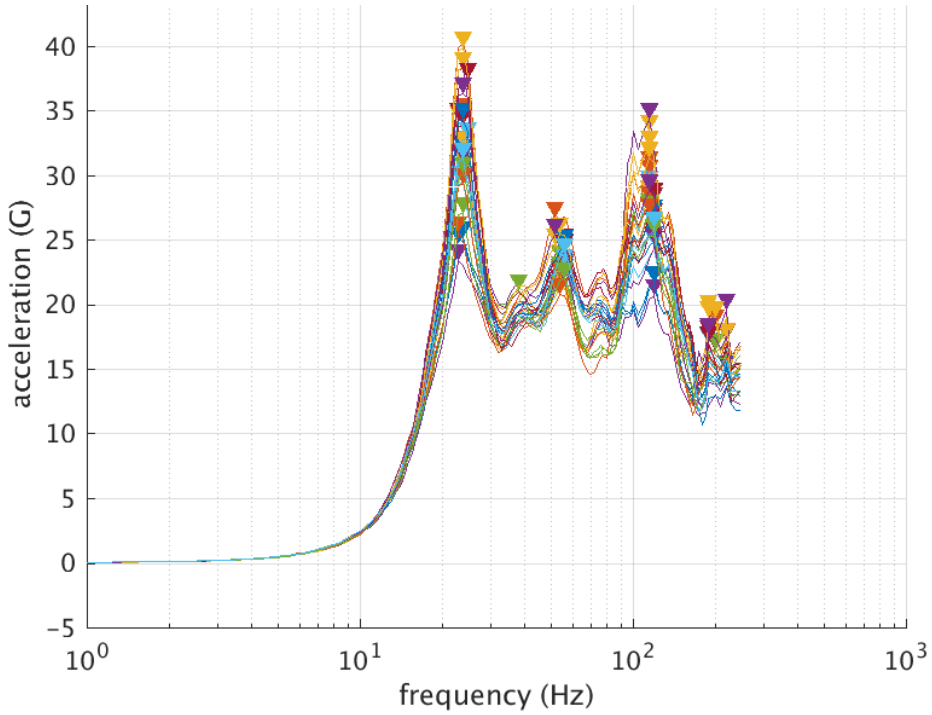

Figure 4.22 Acceleration Response Spectra Overlay of RVCH Lateral Motion

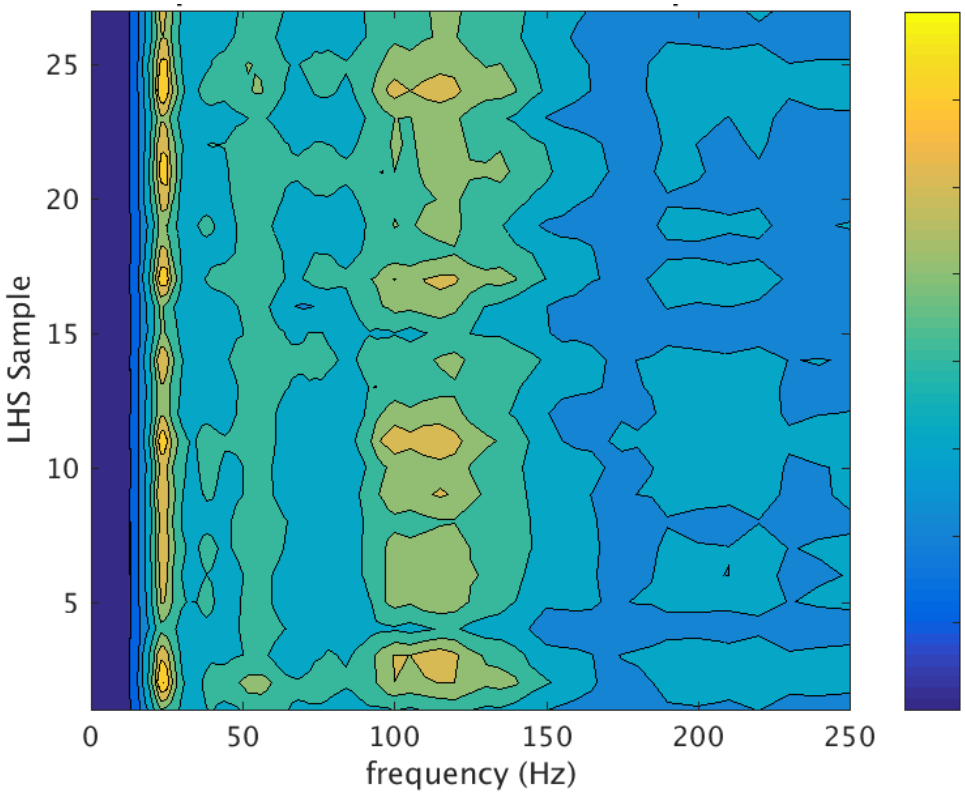

Figure 4.23 Contour Plot of Acceleration Response Spectra (acceleration magnitude represented by color scale)

Maintaining parameter variations of ±5% (as used in Sections 4.2.1 and 4.2.2), and focusing on the acceleration magnitude of the first resonant frequency as the output of interest, Figure 4.24 shows the first-order global sensitivities (computed via FAST per Section 2.2.1) from the full-order SFEM. Comparing these sensitivities to those for the maximum acceleration over time per Figure 4.8 reveals that the importance of parameters ζ and N_6 is significantly increased while k_9 is decreased, and the remaining parameters are of near-negligible importance which is the same for both examining the maximum acceleration over time and the spectral acceleration. The change in relative importance between the ζ, N_6, and k_9 terms likely has some to do with the agreement between resonant frequencies of the associated components (i.e., the specific RVI components at location 6 or 9) and the frequency at which the maximum acceleration occurs. That is, per Figure 4.22 and Figure 4.23, the maximum acceleration occurs at approximately 25 Hz, and so it may be that the resonant frequency of sub-components near location 6 (which pertain to N_6) fall closer to 25 Hz than those at location 9 (which pertain to k_9), and so variance in N_6 imparts greater variance in the acceleration response at that frequency than does variance in k_9. Furthermore, the vibration amplitude is known to vary in inverse proportion of the square root of the damping ratio ζ (Blevins, 2001), and so this parameter would likewise be expected to strongly influence the magnitude of the spectral response.

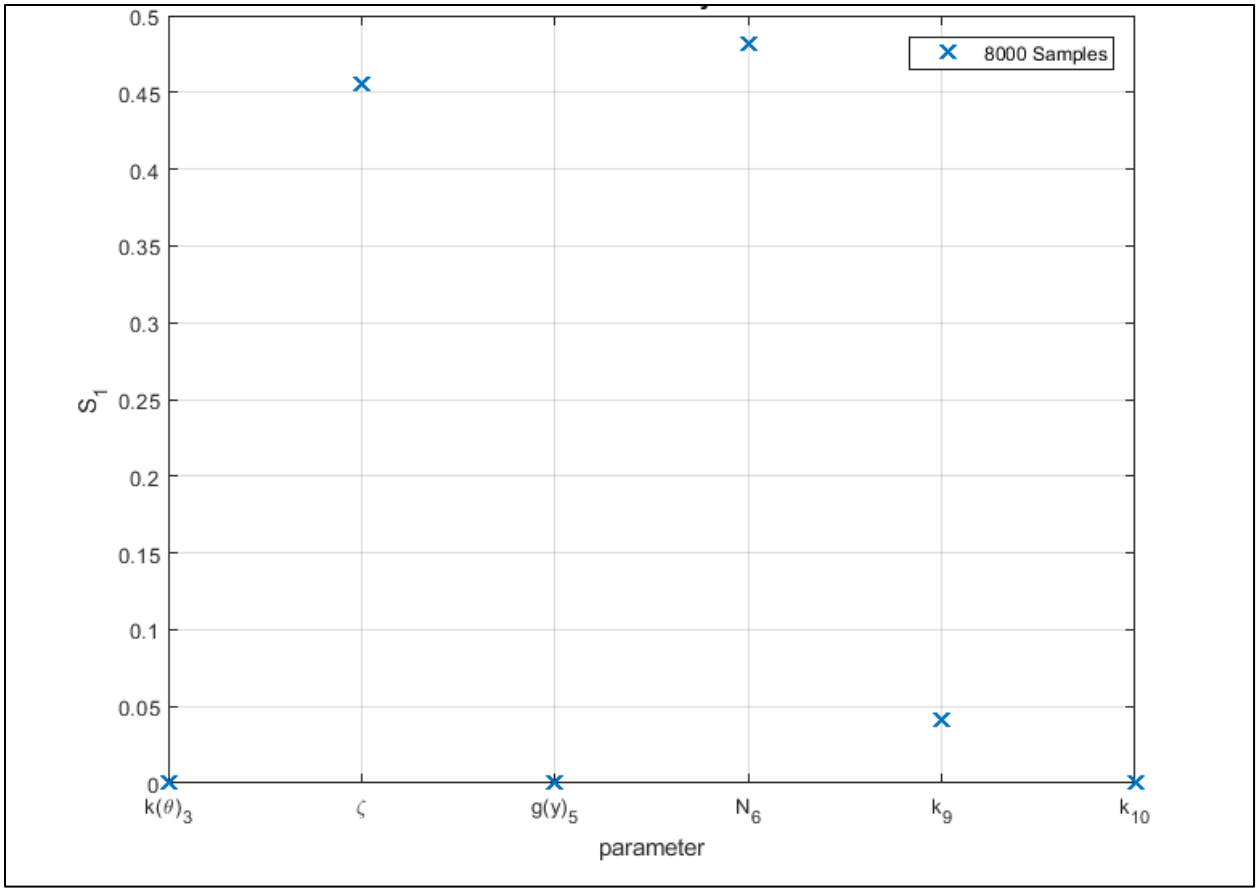

Figure 4.24 Global Sensitivity Indices for System Finite Element Model Response Spectra

Figure 4.25 shows convergence of the first-order sensitivities when Kriging surrogates are trained with LHS and LPSS for a series of training set sizes. It is apparent that at low sample sizes of 27 and 64, the GSA error associated with Kriging surrogates trained by LPSS is lower than the error for those trained by LHS, in some cases by over an order of magnitude. The improvement offered by LPSS makes some intuitive sense because the frequency, ω, which may be simply viewed as a function of $\sqrt{\frac{k}{m}}$, would be directly affected by parameters that affect the mass or stiffness (i.e., five of the six parameters studied here), and thus establishing subdomains in which such terms are grouped according to their interactions via LPSS could result in lower errors as compared to LHS which does not establish such subdomains. Correspondingly, the range of variation in the spectral acceleration being greater than the range of variation in the maximum acceleration of the transient could further serve to render LPSS more effective in yielding lower errors at small sample sizes.

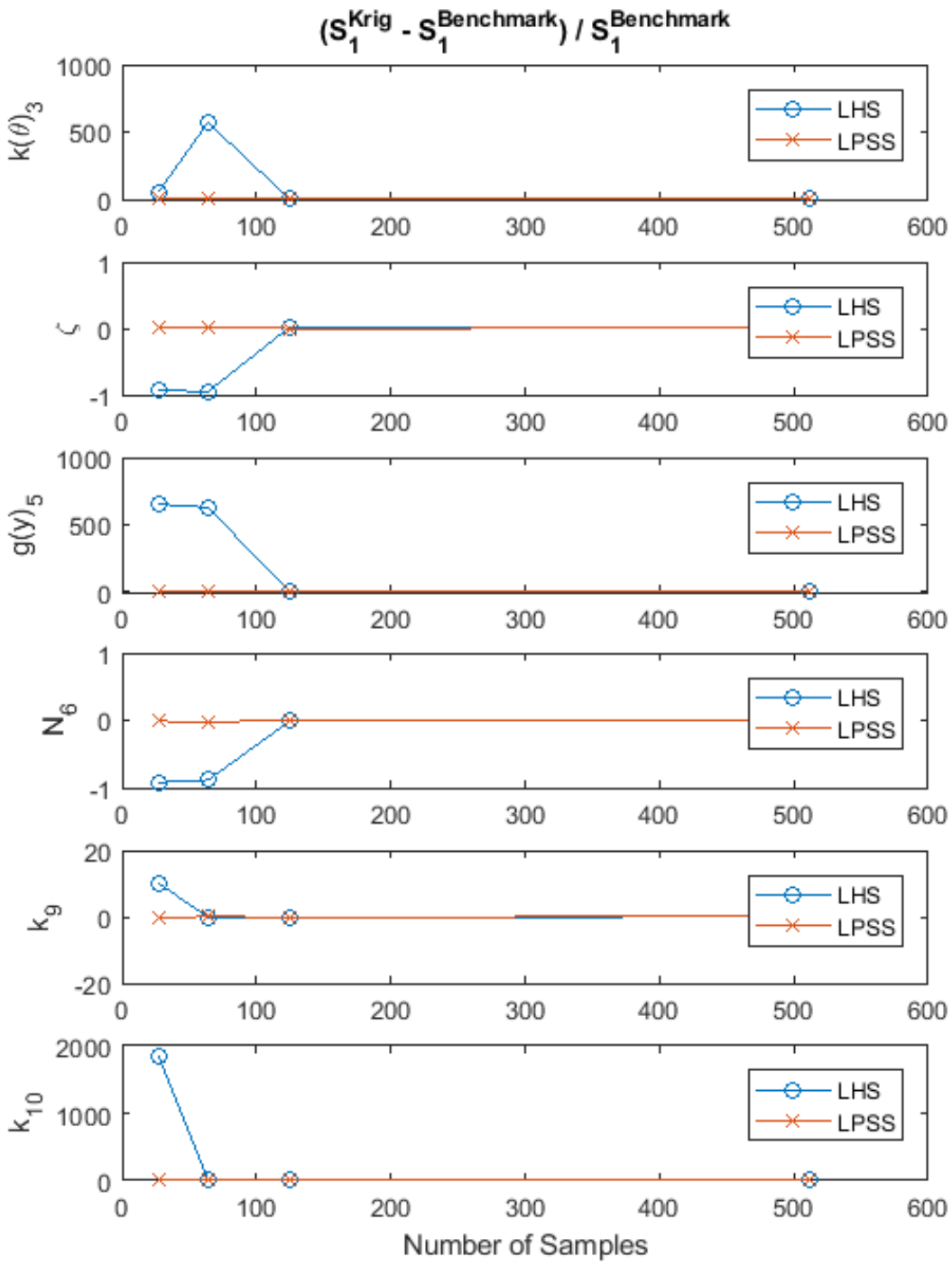

Figure 4.25 Convergence of Surrogate-based Global Sensitivity Indices for System Finite Element Model Acceleration Response Spectra

4.2.4 Surrogate Model Accuracy

Using the Relative RMS and Maximum Relative Residual error metrics per Section 2.3.3 for the SFEM surrogate models, Figure 4.26 illustrates the surrogate model accuracy with respect to an increasing number of samples for the case of the relatively small ±5% range of variation on the input parameters. Consistent with prior observations from Chapter 3.0 for the UI FIV model, as well as the lower GSA error for SFEM from Section 4.2.2, the surrogate model verification error associated with LPSS is generally between 10 – 25% lower than that for LHS at the lower sample sizes of 27 and 64.

Figure 4.27 shows the surrogate model accuracy with respect to sample size, but with a larger ±25% range of variation on the same input parameters. The increase in surrogate model accuracy with LPSS compared to LHS is even more significant than for the larger range of parameter variation. For example, at 27 samples the surrogate verification error is approximately 50% lower for LPSS than for LHS. The difference between the ±5% and ±25% case is consistent with the interaction effects also being more significant with a larger range of variation, such as was illustrated in Figure 4.15 and Figure 4.16. Such consistency is considered reasonable as the stronger interaction effects manifest themselves in effective variance reduction provided by LPSS, and therefore a greater degree to which surrogate model accuracy is improved for smaller sample sizes, when compared with LHS.

Even at the lowest sample size of 27, the surrogate model verification error for LPSS was less than 1.5% and 3% for the Relative RMS and Maximum Relative Residual error metrics, respectively. For most practical purposes, these errors are small enough to render such a surrogate acceptable, which suggests that a mere 27 sample computational DOE can prove quite useful for this application.

Of additional interest is the output response histogram shown in Figure 4.28 for three different numbers of LHS samples. It may be seen that increasing the sample size does introduce variations in the distribution of results, but the magnitude of such variations is judged negligibly small. This variation in the output is consistent with the surrogate model errors in this range of sample sizes presented in subsequent subsections.

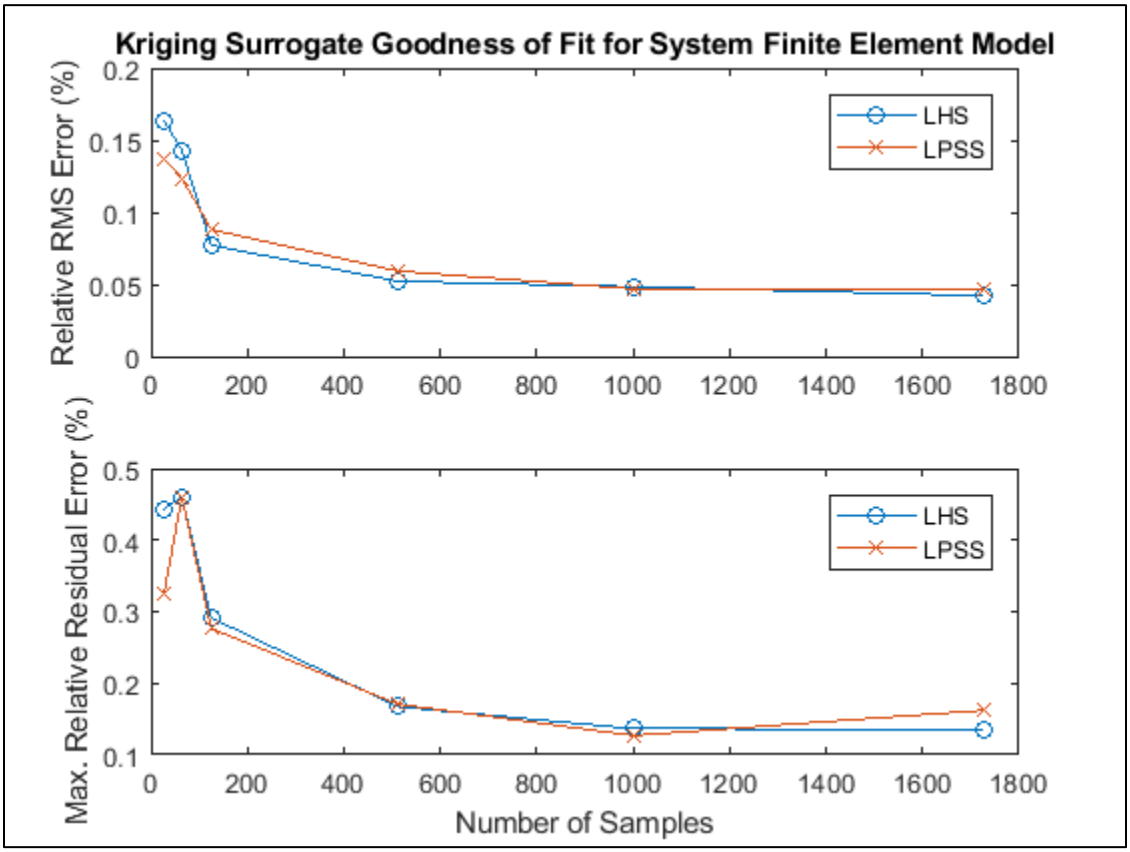

Figure 4.26 Verification of Kriging Surrogate for SFEM with ±5% Input Parameter Variation

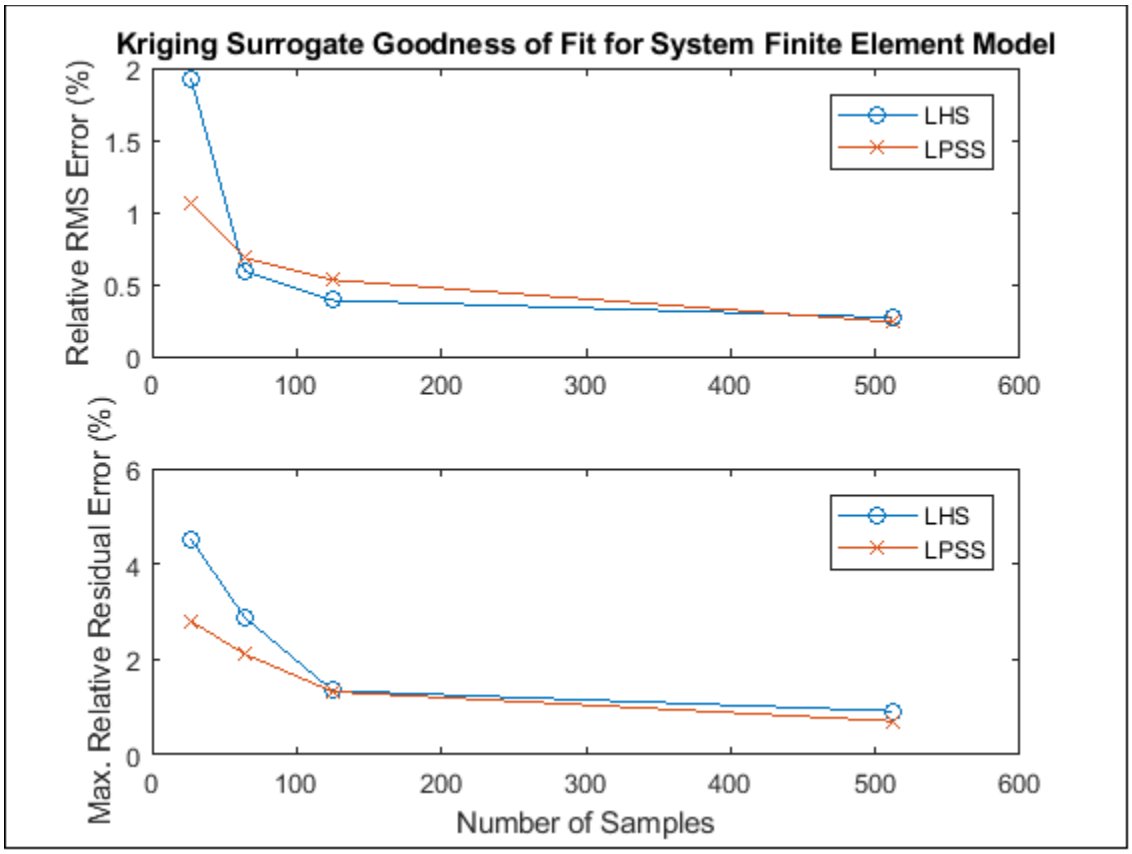

Figure 4.27 Verification of Kriging Surrogate for SFEM with ±25% Input Parameter Variation

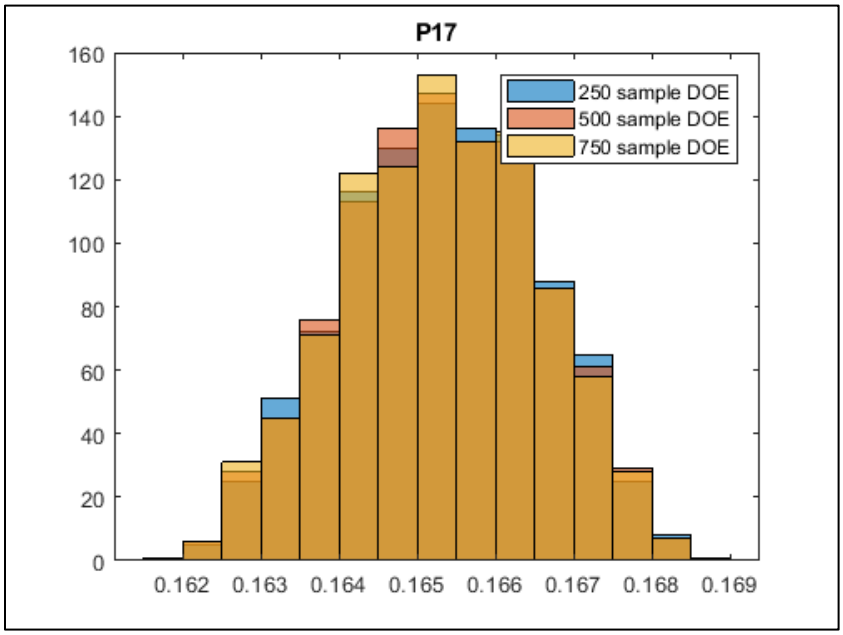

Figure 4.28 Histogram of SFEM Output P17 with Varying Number of Training Samples

4.3 Conclusions from Evaluation of System Finite Element Model

This chapter demonstrates the extensibility of the surrogate-based GSA methodology shown useful for a reactor subassembly subjected to stationary random vibration in Chapter 3.0 to a reactor equipment SFEM, for which the finite element model included non-linearity and was subjected to non-stationary loading. The SFEM was parameterized with a total of sixteen model parameters for which there are plant-to-plant variations or for which the magnitude changes over the course of the reactor life due to aging. Of these sixteen, the six most dominant parameters were chosen to explore for the GSA studies. As a baseline case, a parameter variation range of ±5% about nominal was chosen to represent the normal amount of variation for these parameters amongst the operating fleet of PWRs, and a larger parameter variation range of ±25% was additionally studied to represent further departure from nominal values that some parameters may experience due to nuclear plant aging-related degradation mechanisms. Interaction effects were observed between those six parameters from a factorial DOE, which were used to determine the arrangement of the subdomains for LPSS.

The LPSS and LHS methods were employed to sample the parameter space from which Kriging surrogates were constructed and used to estimate GSA. The use of LPSS was again shown to be particularly effective at providing accurate surrogate models even for low sample sizes. For the smallest number of samples explored, and using LPSS, the magnitude of the surrogate model verification error was approximately 0.3% Maximum Relative Residual when using ±5% variation about the nominal values of model parameters, and less than 3% when using ±25% about nominal. The corresponding Maximum Relative Residual errors associated with LHS increased to approximately 0.5% and 5%, for the ±5% and ±25% parameter variation cases, respectively. Although LHS provides a less accurate surrogate model at small sample sizes, the

resultant errors may be considered acceptable for many practical engineering applications. Thus, the relatively accurate surrogate models obtained through both LHS and LPSS provides confidence in the stability of the result. Such stability may permit one to not invoke advanced sampling methods if the computational cost is tractable (i.e., generating a larger number of samples) and/or error requirements do not demand optimal accuracy.

However, despite the low surrogate verification errors at low sample sizes, the GSA error for some model parameters approached 25% or over 100%, for the ±5% and ±25% variation cases, respectively, for these same sample sizes. For the studies considered herein, approximately 500 samples were required in order for the GSA error to fall below 5% for all parameters, for both sampling strategies. Therefore, although the effectiveness of LPSS is evident at low sample sizes, the GSA error is much higher than the surrogate verification error, which may suggest a need for much larger sample sizes to achieve accurate global sensitivities. This may be visualized in Figure 4.29 and Figure 4.30 within which the RRMS surrogate verification error, as a percentage, is compared with the absolute value of GSA error (which is not a percentage but rather computed as a ratio of $\frac{S_1^{Krig} - S_1^{Benchmark}}{S_1^{Benchmark}}$). Thus, per Figure 4.29 and Figure 4.30, the rate of improvement of both error metrics with respect to increasing the number of samples is comparable, but the magnitudes differ by approximately a factor of 100.

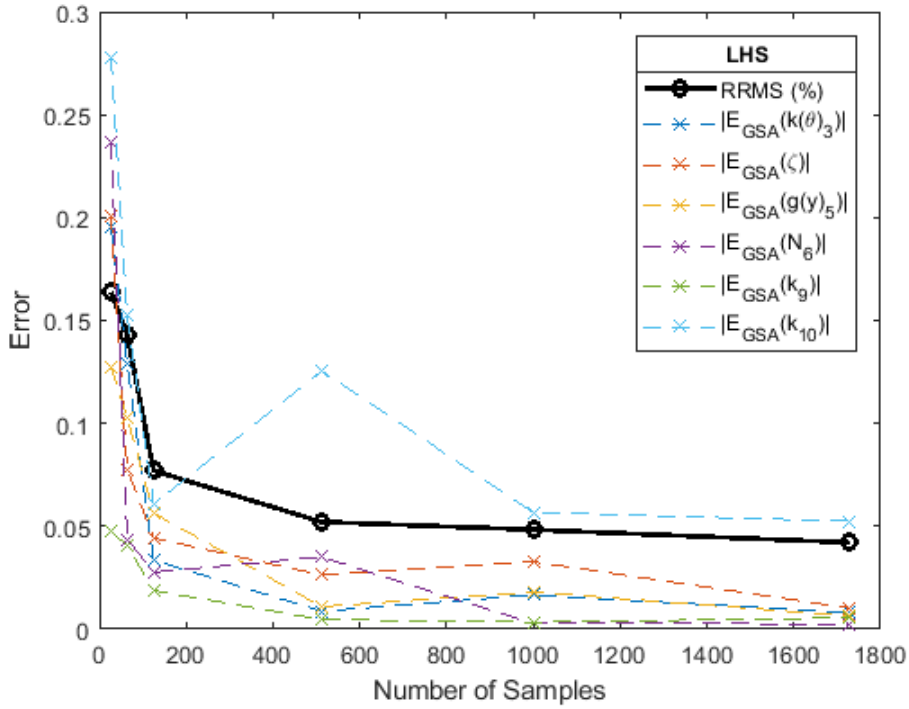

Figure 4.29 Overlay Comparison of Surrogate Verification RRMS Error and GSA Error for LHS

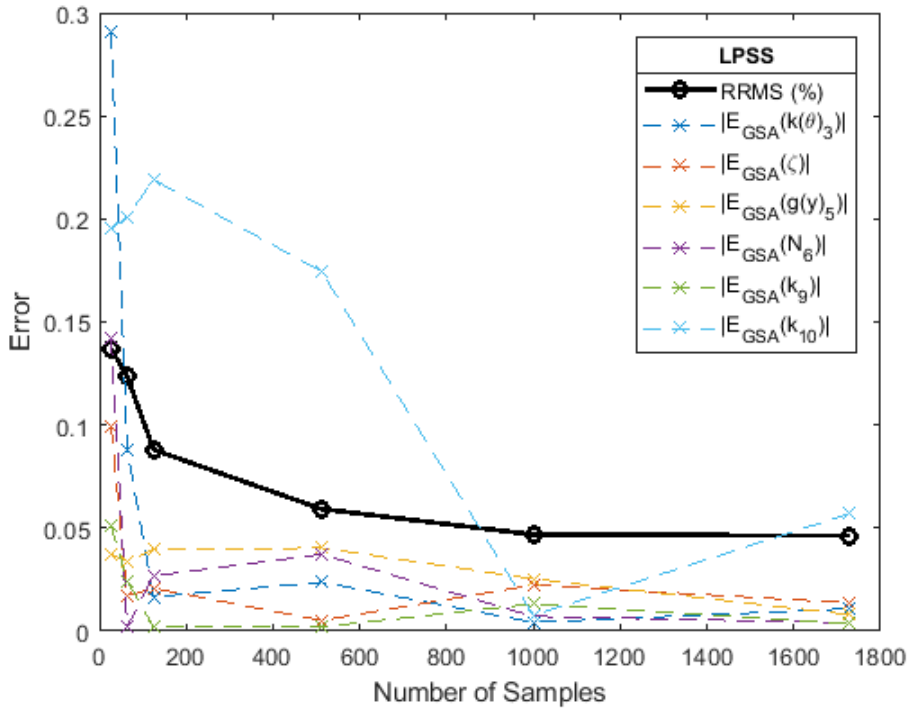

Figure 4.30 Overlay Comparison of Surrogate Verification RRMS Error and GSA Error for LPSS

In such a case that a larger sample size is required, at least for this application, it appears that the differences imparted by LHS or LPSS are small. That is, upon examination of Figure 4.10 and Figure 4.11, for example, the GSA errors are of similar magnitude between the LHS and LPSS methods. Thus, to achieve small GSA errors, a larger number of samples are required, which renders the improvements offered by LPSS over LHS somewhat less apparent. The importance of the relationship between verification and GSA error should be emphasized for the engineering practitioner; just because the surrogate verification error is negligibly small, that does not necessarily indicate that the associated GSA error will be similarly small.

Per the general topic of numerical coarsening, as explored in many domains such as meshing of rigid objects in finite element analysis (Hattangady, 1999) and data-driven finite element modeling (Chen, et al., 2015), derivatives tend to coarsen a model (higher error), while integrals tend to smooth (lower error). In that sense, since sensitivity analyses may be thought of as taking derivatives across a high-dimensional space, it is postulated that the expansion of error from surrogate verification to GSA may be likened to a form of numerical coarsening.

5.0 Effectiveness of Surrogate Modeling Methodological Changes

Three methodological changes, and their effectiveness for direct surrogate modeling and subsequent GSA of the SFEM addressed in Chapter 4.0 are explored in this chapter. First, whereas the studies in both Chapter 3.0 and 4.0 employed Kriging surrogates of a single mean and variation kernel, Section 5.1 explores the effectiveness of an aggregate surrogate modeling method termed Genetic Aggregation per Section 2.3.2. Second, whereas the studies in both Chapter 3.0 and 4.0 used sampling methods (e.g., FAST, LHS, and LPSS) which established the sampling points prior to Surrogate training, an adaptive sampling approach is presented and likewise deployed in Section 5.2 per Section 2.4.3. Such an approach adaptively samples the parameter space based upon the surrogate verification error quantified during the training procedure and is thus intended to provide a means of intelligently selecting samples to more effectively minimize error as compared to non-adaptive sampling approaches. Finally, the different insights offered by variance-based GSA via FAST (Section 2.2.1) and rank correlation-based GSA (Section 2.2.2) are shown in Section 5.3, in order to emphasize to the practitioner the careful interpretation required in considering different types of GSA results.

5.1 Effect of Aggregate Surrogate Model

As in Chapter 4.0 , relative RMS and maximum relative residual surrogate verification quantities, as described in Section 2.3.3, are used as the error metrics of interest in this section. The main idea behind the use of genetic aggregation is that there are a large number of combinations for possible surrogate models. The algorithm thus helps exploring the set of surrogate models since it is impractical to compute all possible surrogates. It is recognized that numerous techniques of building aggregate surrogate models could be explored, each with certain advantages and disadvantages, but this brief study is intended to show the potential benefit offered by the use of such advanced approaches to surrogate modeling for this application.

For equivalent computational DOEs, the surrogate model error between Kriging per Section 2.3.1 (with ordinary trend and ellipsoidal correlation) and genetic aggregation surrogate per Section 2.3.2 is compared in Table 5.1 and Table 5.2 for LPSS and LHS, respectively. Irrespective of sample size, genetic aggregation consistently provides a significant reduction in both error metrics, for both LHS and LPSS. Specifically, the relative RMS error decreases by between 5 – 20% for three different sample sizes.

Considering LHS with 27 samples to be a "base case", it may be seen that the reduction in error is more substantial by using LPSS to train an equivalent Kriging surrogate model type than maintaining the LHS DOE, but changing to an aggregate surrogate model. That is, the relative RMS error reduces from 1.29% for LHS to 1.07% for LPSS when using a Kriging surrogate, whereas the RMS error only reduces to 1.21% when maintaining the LHS DOE but switching from a single kernel Kriging surrogate to a genetic aggregation surrogate. At the higher sample sizes of 64 and 125, LPSS generally shows higher relative RMS error but lower

maximum relative residual error than LHS. Considering the respective error definitions in Table 2.2, this relationship suggests that the maximum amount by which the surrogate prediction differs from the full-order model solution is best minimized by LPSS, while an averaged amount by which many surrogate predictions differ from corresponding full-order model solutions is effectively minimized by LHS.

Table 5.1 Comparison of Surrogate Error Metrics between Kriging and Aggregate Surrogate Model using LPSS

Sample Size	Kriging		Genetic Aggregation	
	Relative RMS (%)	Max. Relative Residual (%)	Relative RMS (%)	Max. Relative Residual (%)
27	1.07	2.80	1.01	2.47
64	0.68	2.11	0.60	1.67
125	0.53	1.32	0.42	0.89

Table 5.2 Comparison of Surrogate Error Metrics between Kriging and Aggregate Surrogate Model using LHS

Sample Size	Kriging		Genetic Aggregation	
	Relative RMS (%)	Max. Relative Residual (%)	Relative RMS (%)	Max. Relative Residual (%)
27	1.29	3.31	1.21	3.94
64	0.59	2.89	0.49	1.57
125	0.39	1.34	0.33	1.20

Examining the sequential generations, as governed by a generic algorithm per Equation (46), which aim to develop a weighted combination of surrogates which best minimize the penalized predictive score (PPS) for output parameter P17 in Table 5.3, it may be seen that after the 2nd generation of an aggregate surrogate model, no additional changes were made between the 3rd and 12th generation. Specifically, after 2 generations the genetic aggregation found an

optimally-weighed ensemble of seven surrogate models consisting of three Kriging surrogates, three support vector regression surrogates, and one moving least squares surrogate. In that sense, the RMS error obtained by genetic aggregation algorithm converged because after only one generation, the optimal combination of surrogate models (i.e., those which best minimized the PPS error) was found and ceased to change upon subsequent generations of the genetic algorithm. Considering another output parameter P18, the formulation of aggregate surrogate models corresponding to twelve generations of the genetic aggregation method are shown in Table 5.4, from which it may be seen that the optimal combination of surrogate models changed three times between the first and twelfth generation, with reduction in RMS error upon each change. Thus, in the case of output parameter P18 the optimal aggregate model consisted of three different Kriging models of differing kernel variation and regression functions. The choice to limit the number of generations to twelve was somewhat arbitrary and based on visual inspection of the error reduction achieved by increasing the number of generations.

Table 5.3 Genetic Aggregation Training Generations for output parameter P17

Generation	Weight	Surrogate Type	RMS Error (ϵ)
1	1.0	Kriging, Gaussian kernel, anisotropic variation, full quadratic regression	0.000473
2 – 12	0.184	Linear basis function, full quadratic regression, Gaussian support vector kernel type	0.000441
	0.026	Support vector regression, $\epsilon = 1 \times 10^{-8}$, Gaussian support vector kernel type, Laplace loss function	
	0.190	Kriging, Gaussian kernel with anisotropic variation, full quadratic regression	
	0.146	Kriging, damped Sin kernel with anisotropic variation, constant regression	
	0.132	Moving least squares, Wendland weighting function type, full quadratic polynomial type	
	0.188	Kriging, Bessel kernel with anisotropic variation, full quadratic regression	
	0.134	Linear basis function, pure quadratic regression, damped Sinus support vector kernel type	

Table 5.4 Genetic Aggregation Training Generations for output parameter P18

Generation	Weight	Surrogate Type	RMS Error (ϵ)
1	1.0	Kriging, damped Sin kernel with anisotropic variation, constant regression	0.000732
2 – 9	0.143	Support vector regression, $\epsilon = 1 \times 10^{-8}$, Gaussian support vector kernel type, Laplace loss function	0.000659
	0.857	Kriging, damped Sin kernel with anisotropic variation, constant regression	
10	1.0	Kriging, cubic kernel with anisotropic variation, constant regression	0.000622
11 – 12	0.400	Kriging, cubic kernel with anisotropic variation, constant regression	0.000609
	0.330	Kriging, damped Sin kernel with anisotropic variation, constant regression	
	0.270	Kriging, cubic kernel with anisotropic variation, linear regression	

5.1.1 Correlation of Cross-Validation and Verification Points

The goodness of fit metrics shown in Section 2.3.3 were calculated for SFEM in three different ways:

1. **Training Points**, which measure the quality of the interpolation. These are the points provided to fit the surrogate model parameters. It is expected that the points with which the Surrogate model is trained may be accurately reproduced by the Surrogate model itself. Therefore, good quality is always expected for the learning points.
2. **Verification Points**, which measure the quality of the prediction against an independent set of data separate from the training data. This is considered the standard and intuitive way by which Surrogate model quality may be assessed.
3. **Cross-Validation Points**, which measure the stability and reliability of the Surrogate model by leaving out a certain portion of the training set, for use in testing against the model built from the samples included in the training set, and is repeated multiple times (e.g., k-fold or leave-one-out (LOO) methods).

The error with respect to the verification and cross-validation points was very well-correlated, as shown in Figure 5.1, with which the verification point error may be partially inferred from the cross-validation point error. Within this plot, each circle represents the error for a given SFEM output which was predicted from the surrogate, for a total of 26 predicted quantities. In most of the studies in Chapters 3.0 and 4.0 , in addition to the training samples, 100 additional samples were generated from the full-order model for purposes of computing error for a set of verification points. Therefore, the cross-validation error provides a reasonable means of assessing the model accuracy. The strong degree to which the circles fall along the 45° line illustrated by Figure 5.1 suggests that, for this application, the cross-validation error provides an accurate estimate of the verification error. In such a case, the need to compute additional samples for verification purposes could be alleviated. Furthermore, Figure 5.2 illustrates a histogram of the percent differences accumulated across multiple plants (i.e., multiple sets of data corresponding to that shown in Figure 5.1 for one plant). From Figure 5.2 it

may be seen that the percent differences are centered about zero, although in many cases the cross-validation error was very low with respect to the verification point error as evidence by the large number of data points with -100% difference. Thus, the use of errors for the cross-validation points can prove insightful for assessing surrogate model quality for the SFEM analysis. Although the SFEM explored herein was not particularly computationally expensive to execute, such an observation could prove quite useful for applications with substantially more expensive full-order models.

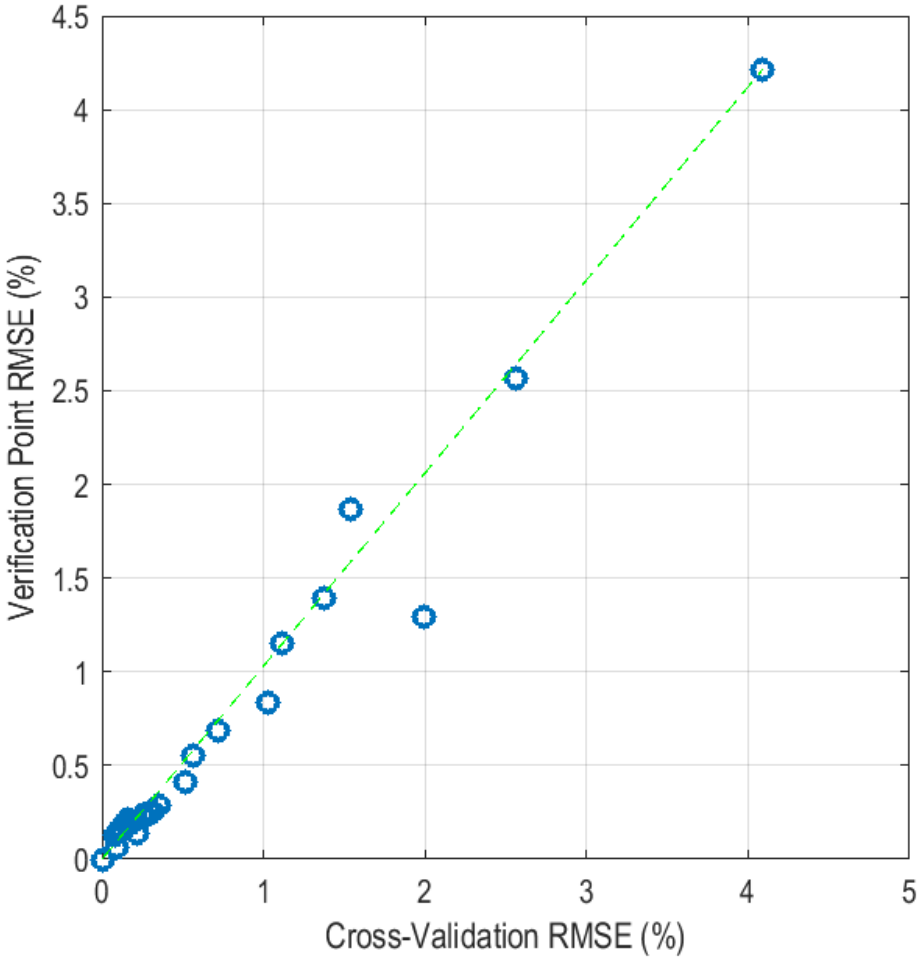

Figure 5.1 RMS Error Metric Comparison between Verification and Cross-Validation Points

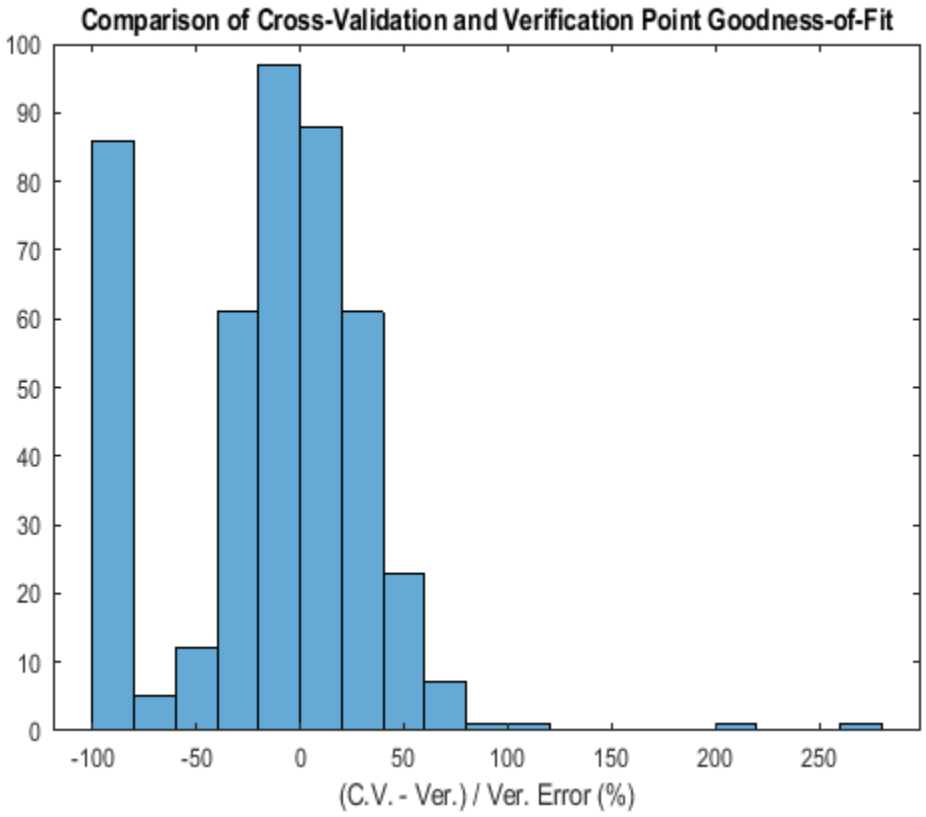

Figure 5.2 Histogram of Percent Differences between Cross-Validation and Verification Point Errors

5.2 Effect of Adaptive Sampling

This subsection is intended to offer some preliminary insight as to the potential improvements offered by adaptively sampling the parameter space, as opposed to treating the parameter sampling, surrogate training, and surrogate verification as a serial process.

The case in which 125 LHS samples were used to train a Kriging surrogate for SFEM with ±25% parameter variation was chosen for study in this subsection, considering two

additional cases of varying the balance of initially-sampled versus adaptively-sampled points as described in Table 5.5, and the same surrogate verification error metrics previously considered for SFEM. It may be seen that as the relative proportion of adaptive samples is increased, the RMS and relative residual errors with respect to the verification points decreases, although the cross-validation error actually increased. Thus, the differences between verification and cross-validation error observed in this introductory study of adaptive sampling suggests that the correlation between these error metrics may not be as strong as suggested by the study in Section 5.1.1.

Table 5.5 Cases Considered to Study the Effect of Adaptive Sampling

Total Number of Samples	N_{LHS}	$N_{Adaptive}$	Cross-Validation		Verification	
			RRMS	MRR	RRMS	MRR
125	125	0	0.209	0.627	0.331	1.197
	75	50	0.248	0.709	0.251	1.067
	50	75	0.236	0.730	0.217	0.588

To provide some graphical insight as to the adaptive refinement points which were selected subsequent to LHS, based on the UP-distribution described in Section 2.4.3, Figure 5.3 through Figure 5.5 show the points which were sampled for the 3rd case shown in Table 5.5 (50 LHS, 75 adaptive). Specifically, Figure 5.3 and Figure 5.4 show the scatter of points selected across the domain of parameters P12 (N_6) and P15 (k_9), which are the two parameters which imparted the greatest sensitivity to RVCH acceleration of the SFEM (per Figure 4.8). When viewed with respect to a single parameter, it may be observed that the majority of the adaptive refinement points were placed towards the extreme ends of the parameter space. Extending this to view with respect to two parameters simultaneously, Figure 5.5 shows the scatter of points

selected across the bivariate parameter space including both P12 and P15 (i.e., Figure 5.5 shows the combination of data points from Figure 5.3 and Figure 5.4). While upon visual inspection, a large proportion of points within the parameter space are placed at the extreme ends (i.e., near a value of either 0.75 or 1.25 for P12 or P15), in actuality the majority of the samples included at least one point selected from an intermediate value for a given parameter, even though this cannot be easily viewed from a scatter plot. The adaptive refinement algorithm used tended to define samples in which four or five of the six parameter values were placed at the extreme, but the remaining one or two parameter values were selected at an intermediate point within the parameter space. The first 20 refinement samples chosen for this case are detailed in Table 5.6, within which the parameter values selected at intermediate points within the parameter space are highlighted yellow. Of the total of 75 refinement samples selected for this study, 50 included a parameter magnitude of intermediate value for one parameter, 32 included an intermediate parameter value for two parameters, 17 included an intermediate parameter value for three parameters, and seven included an intermediate parameter value for four parameters. For a given parameter, between 15 and 22 of the total 75 refinement samples included an intermediate point (e.g., for P15, 19 of the 75 samples included a value between 0.751 and 1.249). The selection of these points by the UP-SMART algorithm (per Section 2.4.3) thus suggests that while many points towards the extreme ends of the parameter space tend to increase universal prediction (UP) variance, and so were chosen in order to optimally decrease UP variance, in the majority of adaptive samples at least one intermediate parameter value was selected.

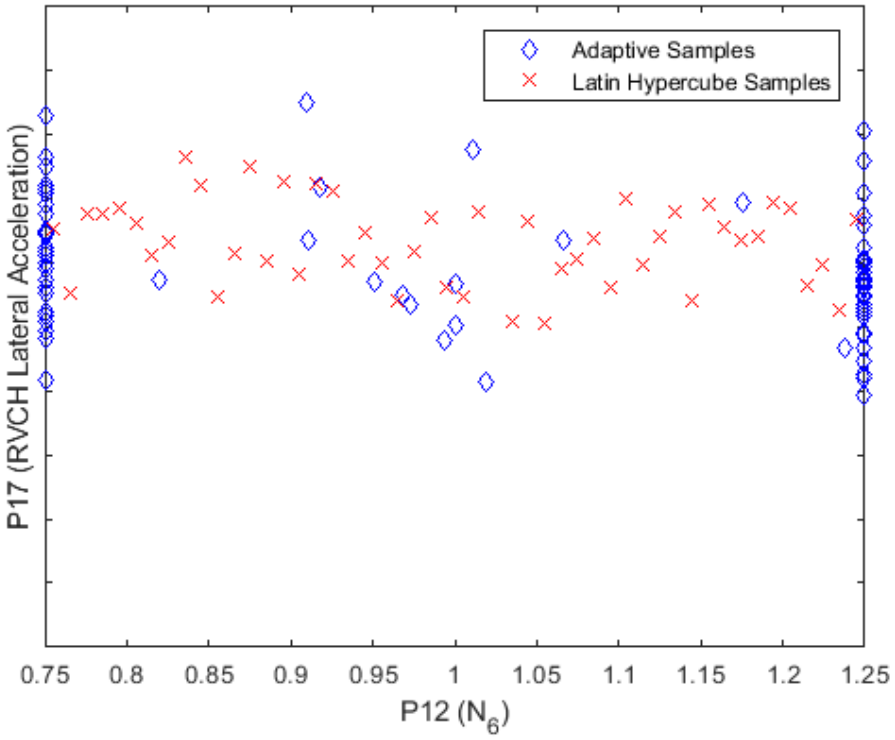

Figure 5.3 Comparison of Latin Hypercube and Adaptive Refinement Samples for P4

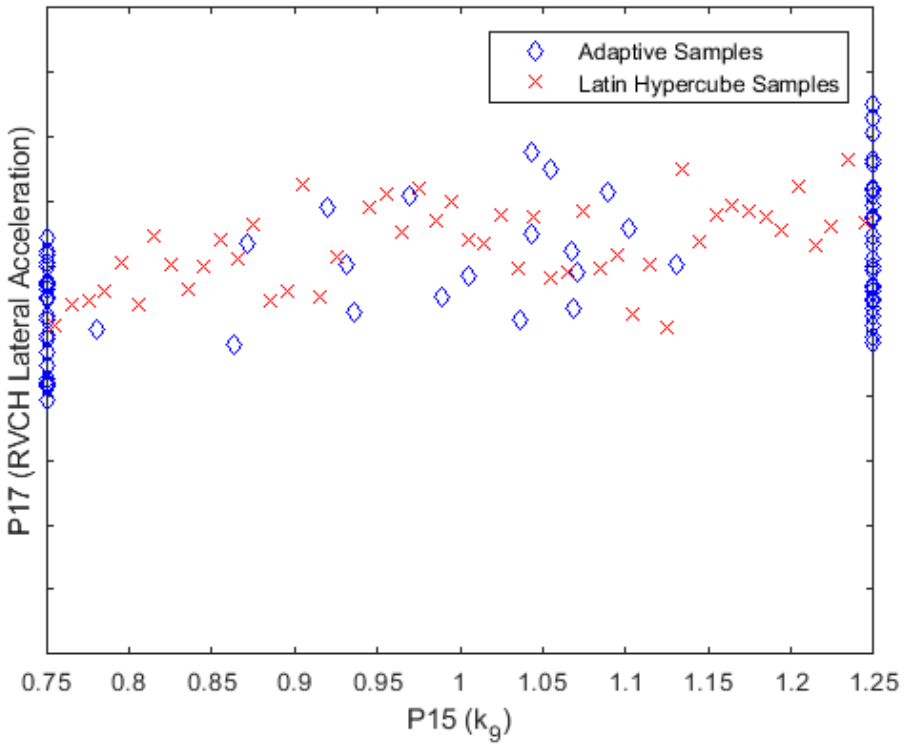

Figure 5.4 Comparison of Latin Hypercube and Adaptive Refinement Samples for P10

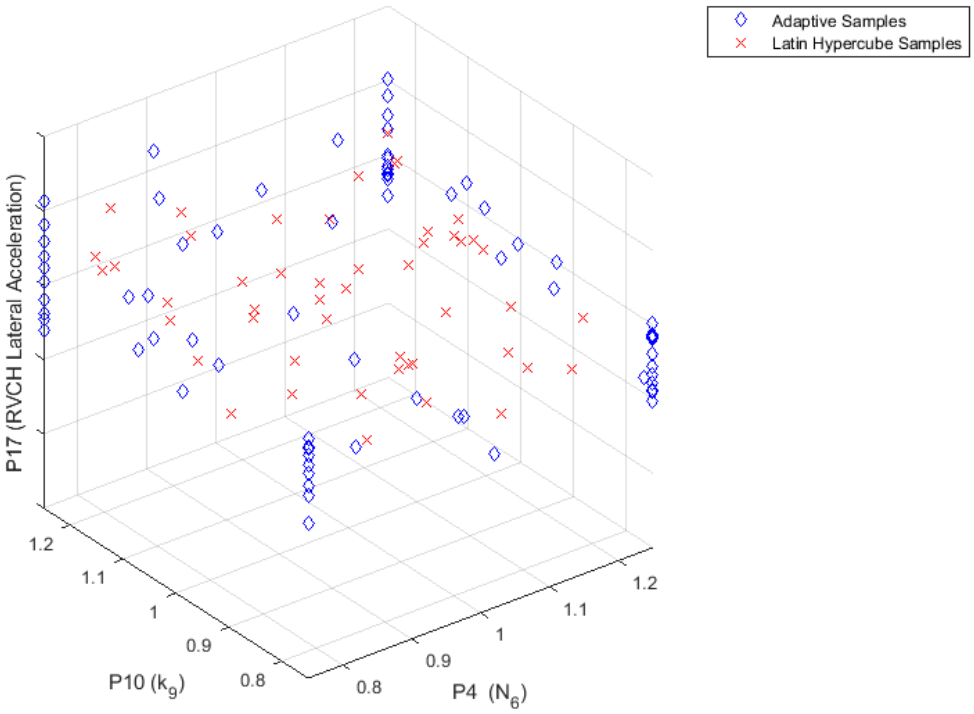

Figure 5.5 Comparison of Latin Hypercube and Adaptive Refinement Samples for P4 and P10

Table 5.6 Adaptive Refinement Point Placement

P4 $k(\theta)_3$	**P10** ζ	**P11** $g(y)_5$	**P12** N_6	**P15** k_9	**P16** k_{10}
1.2495	0.0500	1.2495	1.2495	0.7505	0.7505
1.2495	0.0300	1.2495	1.2495	1.2495	1.2495
0.7505	0.0500	0.7505	1.2006	0.7505	0.7505
1.2495	0.0342	1.2495	0.7505	0.7505	0.7505
0.7505	0.0300	0.7505	1.2495	0.7505	0.7505
1.2495	0.0500	0.7505	1.2495	1.2495	1.2495
0.7505	0.0300	0.7505	0.7505	0.7505	0.7505
0.7505	0.0300	1.0039	1.2495	1.2495	0.7505
0.7505	0.0500	1.2495	1.2495	1.2495	1.2495
0.7505	0.0500	0.7505	1.2156	0.7524	1.2495
0.7505	0.0500	0.7505	0.7505	1.2245	1.2495
0.7505	0.0500	1.2495	1.0097	1.0615	1.2495
1.2495	0.0300	1.2495	1.2495	1.0411	1.2495
0.7505	0.0379	0.7505	1.0036	1.2068	1.2495
1.2495	0.0500	1.2495	0.7505	0.9932	0.7505
0.7505	0.0500	1.2495	0.7505	1.2495	0.8848
1.2495	0.0300	0.7505	0.9154	1.2495	1.1020
1.0014	0.0500	0.7505	0.7505	1.1044	0.7505
1.0033	0.0500	0.9827	1.2495	1.2268	1.2495

5.3 Correlation versus Variance-Based Global Sensitivity Methods

Throughout this dissertation, the FAST method was employed as a variance-based means of computing global sensitivity. FAST computes first-order global sensitivities that equate to those from Sobol's method per (Saltelli, et al., 2008), but Sobol's method can provide additional insight as to the 2nd order sensitivities for situations in which the interaction of multiple parameters significantly imparts variance to key outputs of interest. The need to compute higher-order sensitivities can be deduced by the summation of the first-order global sensitivities.

If the summation of first-order sensitivities is nearly equal to 1.0 (one heuristic would be 0.98 or greater), then the higher order terms are negligible. Indeed, that is the case (and so FAST was clearly appropriate) for all studies performed throughout this dissertation.

While FAST was found to provide useful insight as to relative parameter importance within this work, means of performing variance-based sensitivity analysis (e.g., FAST) is not provided in many commercial software packages. Indeed, numerous commercial software codes provide measures of parameter importance which are based solely on rank correlation, even though some literature recommends computing global sensitivities with multiple methods, since the insights offered by each method may be different. Note that the absolute values of the sensitivities should be considered in comparing relative importance between FAST and rank correlation; the rank correlation provides insight as to if the result is monotonically decreasing or increasing whereas the variance-based global sensitivity does not provide information regarding the directionality of the relationship. With that in mind, a subset of the FAST GSA results are compared to rank correlation based GSA results in this subsection. It is hoped that this subsection may provide some preliminary motivation for engineering practitioners to consider the merits provided by variance-based methods for GSA.

Formulating the problem in an equivalent manner to Section 4.2 using LHS, Figure 5.6 shows the Spearman rank correlation based global sensitivities, from which it may be seen that P11 (vertical gap at location 5, $g(y)_5$ per Table 4.1) is the most influential parameter, followed by P15 (stiffness at location 9, k_9) and P10 (damping, ζ), while P12 (number of components at location 6, N_6) is one of the least influential parameters. The rank correlation measure of relative importance thus provides different results from the variance-based global sensitivities computed via FAST, such as illustrated in Figure 4.8 in which $g(y)_5$ is identified to be among the least

influential parameters and N_6 is among the most influential parameters. To provide additional graphical insight, Figure 5.7 through Figure 5.9 show scatter plots of the Latin hypercube samples used for this study, including a line drawn though the point cloud for which the slope provides some indicative measure of parameter sensitivity. Figure 5.7 and Figure 5.8 show relatively strong and weak negative monotonic trends for P11 and P12, respectively, regarding the relationship between the RVCH acceleration response and the model parameter magnitude, which manifests itself as rank correlation coefficients of approximately -0.6 and -0.2 per Figure 5.6. Correspondingly, Figure 5.9 shows a relatively strong monotonic trend for P15 which manifests itself as a rank correlation coefficient of approximately 0.4 per Figure 5.6. The global sensitivities computed with Spearman's rank correlation are plotted alongside those computed via FAST in Figure 5.10.

As partially illustrated in Figure 4-5 through Figure 4-7, the SFEM results do not necessarily follow a monotonic trend with respect to changes in single model parameters. Therefore, since the rank correlation coefficient is a measure related to monotonicity and the first-order sensitivity computed via FAST is based upon the variance, the differences in relative sensitivity provided by FAST versus rank correlation make some intuitive sense. However, these different measures of sensitivity have implications in terms of characterizing relative importance of parameters in terms of whether the analyst is interested in which parameters best relate to monotonicity of a result, in which case one would use rank correlation, or which parameters best explain the variance of a result, in which case one would use a variance-based method such as FAST or Sobol's method. Therefore, when relative parameter importance is quantified, it is key to describe the method by which that importance was computed so that the

sensitivity analysis result may be appropriately characterized and thus parameters of high importance be intelligently selected.

Furthermore, in subsequent optimization studies which could be informed by global sensitivity results, caution should be used in determining the most appropriate sensitivity method. For example, if one desired to construct an inverse problem in which the gap at location 5 $g(y)_5$ (P11) were of interest to estimate, which is not directly measurable during plant operation, from RVCH acceleration, which is directly measurable during plant operation, then the rank correlation results suggest that a monotonically increasing measurement of RVCH acceleration indicates a correspondingly decreasing $g(y)_5$, although the FAST results suggest that $g(y)_5$ contributes relatively little variance to that measured for RVCH acceleration.

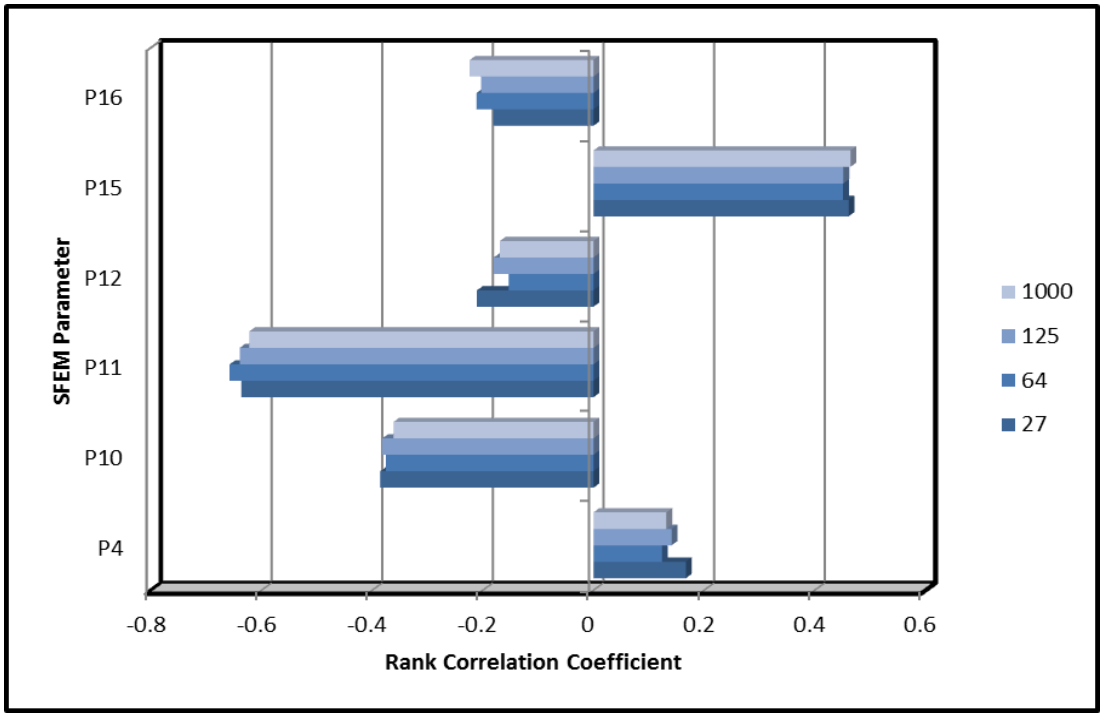

Figure 5.6 Spearman Rank Correlation Coefficient Global Sensitivities for SFEM

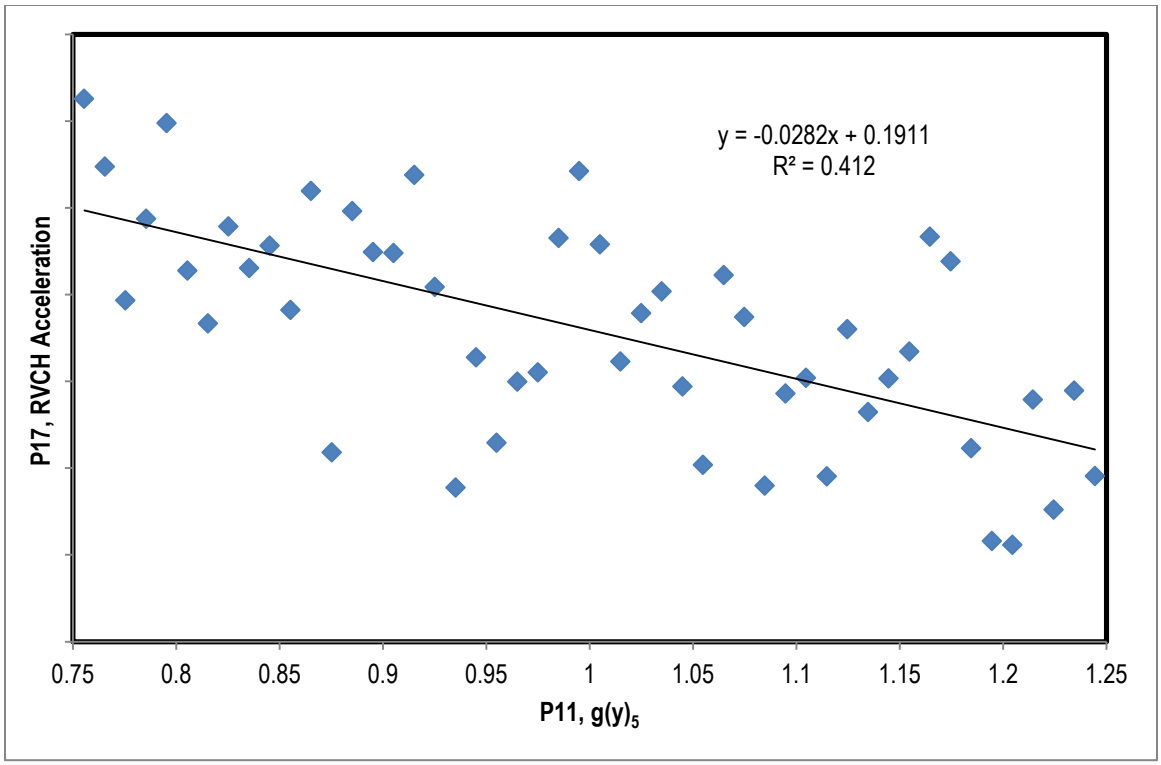

Figure 5.7 Scatter Plot between Vertical Gap at Location 5 (P11) and RVCH Acceleration

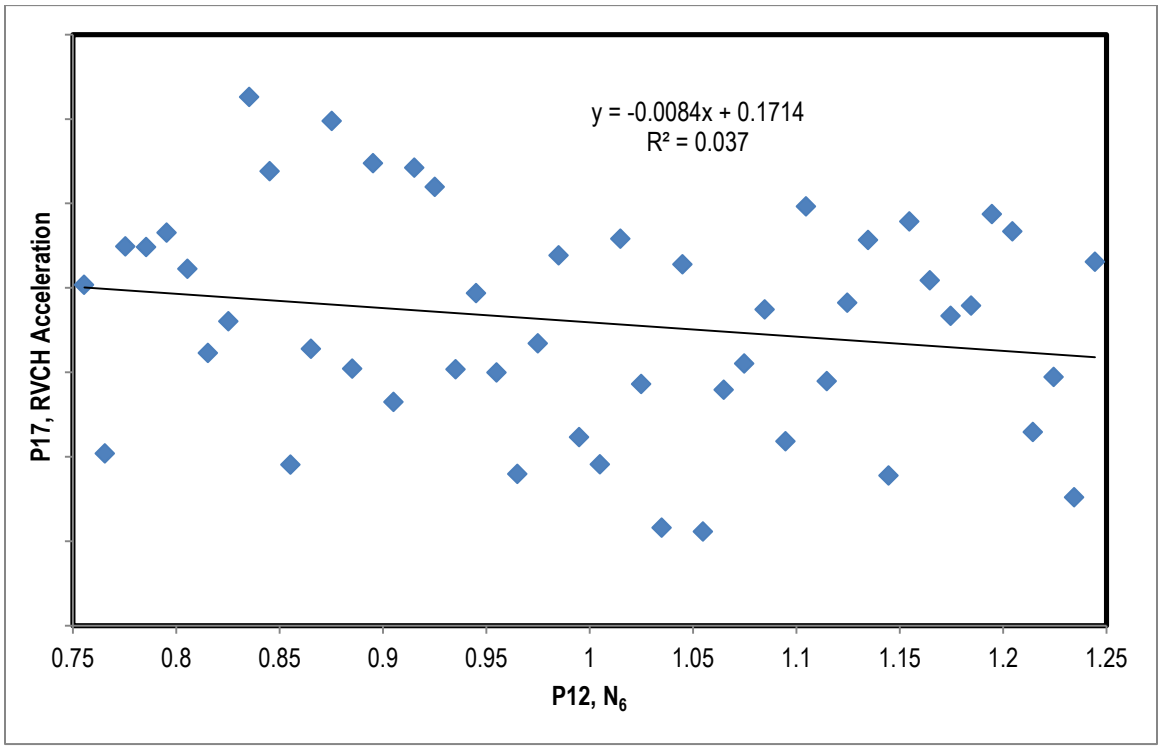

Figure 5.8 Scatter Plot between Number of Components at Location 6 (P12) and RVCH Acceleration

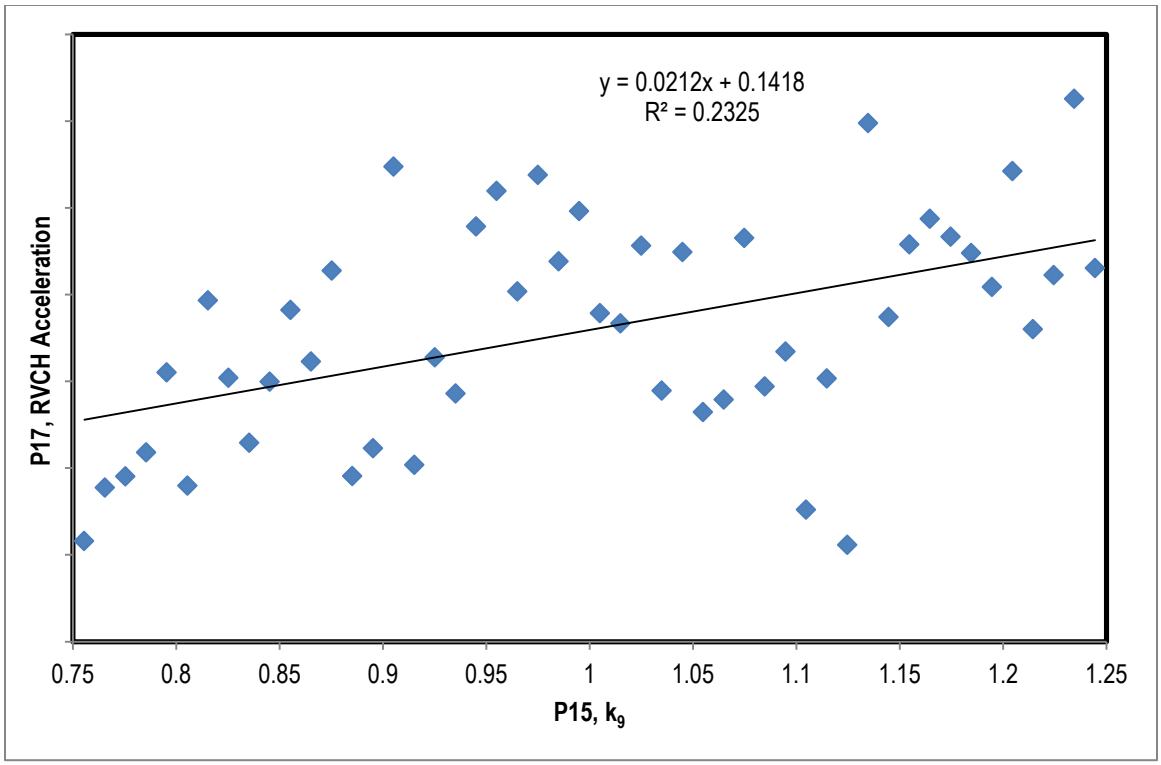

Figure 5.9 Scatter Plot between Stiffness at Location 9 (P15) and RVCH Acceleration

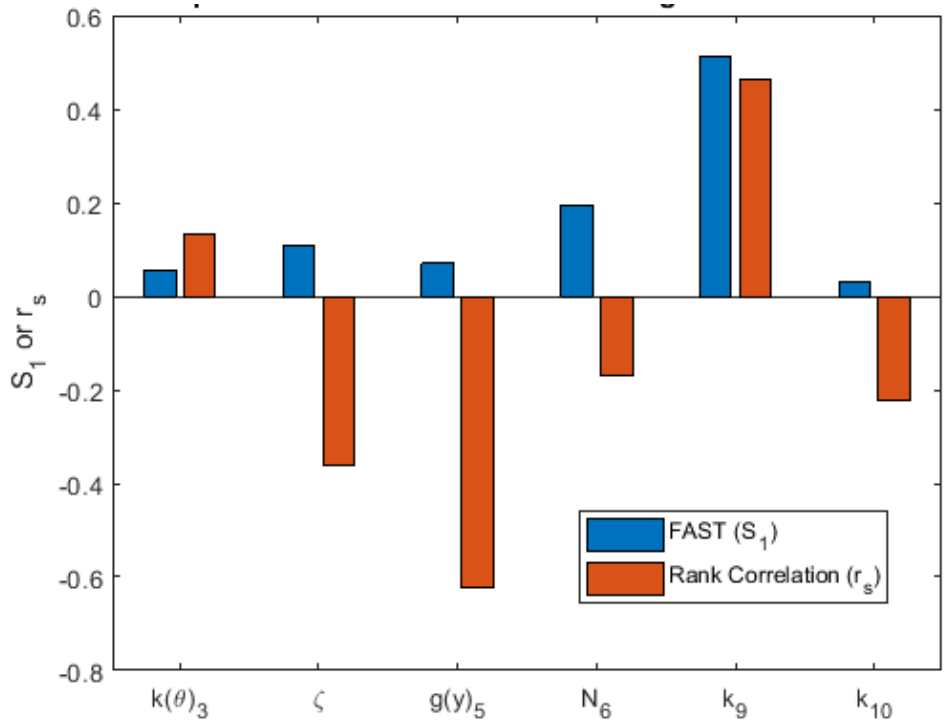

Figure 5.10 Comparison of Global Sensitivities between Variance-based and Correlation Methods

6.0 Conclusions of Dissertation

This work offers a robust perspective as to the pertinence of surrogate modeling and global sensitivity analysis (GSA) for nuclear reactor stochastic dynamics with a focus on structural vibration, as well as the effect imparted through different sampling methodologies for constructing a computational experiment. While direct incorporation of stochastic simulation methods such as surrogate modeling and GSA is presently not commonplace in the context of design and analysis of nuclear power facilities, this dissertation demonstrates the stability of such methods for use in support of two different domains of structural dynamic analysis for reactor assemblies. The agreement in surrogate model accuracy metrics and GSA results amidst multiple sampling techniques provides confidence as to the credibility of such methods.

The computational models selected represent a spectrum of domains of work for practicing engineers in the nuclear industry. The upper internals (UI) flow-induced vibration (FIV) problem studied in Chapter 3.0 examined stationary FIV in order to simulate normal operating conditions experienced on a continuous basis at nuclear power plants, as well as the conditions under which testing is performed prior to starting up a new plant. While the model fidelity of the UI FIV analysis was adequate to resolve strain, the system finite element model (SFEM) analysis studied in Chapters 4.0 and 5.0 constituted a coarser model used for purposes of generating interface loads, which are subsequently used in higher-resolution sub-models for computing strain. The SFEM analysis thus examined the non-stationary loading associated with a loss of coolant accident (LOCA), which is an event necessary to consider in faulted scenarios (i.e., not normal operating conditions) necessary for design basis analysis.

Gaussian Process Regression (Kriging) surrogates were trained by way of Latin hypercube (LHS), Latinized partially stratified (LPSS), and Fourier amplitude sensitivity test (FAST) methods. In the presence of parameter interactions, the effectiveness of LPSS was observed to be most pronounced at small sample sizes in terms of reducing the error with respect to both surrogate model verification metrics as well as the error associated with surrogate-based GSA results relative to those computed directly from full-order models. For the upper internals FIV analysis, both LHS and LPSS were shown to be superior sampling strategies as compared to using FAST directly to sample the parameter space, with respect to training accurate surrogate models to achieve accurate GSA results with a minimal number of samples. For sample sizes greater than 125, nearly equivalent GSA results were achieved for all sampling methods, which demonstrate stability in the surrogate-based GSA methodology, and thus demonstrate the veracity of the GSA results overall, provided sufficient samples can be generated.

For the SFEM analysis, a number of comparison studies were performed. First, Section 4.2.3.1 examines the effect of increasing input parameter variance. The increased parameter variance significantly increased the surrogate model verification error for a given number of samples. The increased variance also tended to exacerbate the parameter interaction effects thus rendering the improvements offered by LPSS over LHS more pronounced, in terms of surrogate verification error. Second, Section 4.2.3.2 examined the effect of changing the output of interest from the maximum acceleration of the time-history to the maximum acceleration at the first mode of the acceleration response spectrum. As the variation in amplitude of vibration at the first mode proved significantly greater than the variation in the maximum amplitude of vibration over time, the improvements offered by LPSS over LHS in achieving accurate GSA results were demonstrated.

Methodological changes are introduced in Chapter 5.0 in order to motivate further research and alert practitioners to the possibilities offered by such changes. Namely, the effect of using an aggregate surrogate modeling approach and effect of adaptive sampling are explored in Sections 5.1 and 5.2, respectively, wherein incremental improvements in surrogate model accuracy by the use of such methods are demonstrated. Also, variance-based GSA versus rank-correlation analysis methods are briefly explored in Section 5.3, wherein the different insights offered by these approaches to quantifying parameter importance are elucidated. Deeper exploration of such methods and variants thereof could constitute a form of future research.

While related work has explored the extensibility of similar principles to those addressed in this dissertation for thermal transient analysis (Banyay, et al., 2012) and hydrodynamics (Banyay, et al., 2014) within reactor applications, it is expected that modeling and simulation of many physics besides structural dynamics could likewise realize the benefits of such methods. Examples could include, but are not limited the simulation of severe accidents for reactor safety analysis and probabilistic risk assessment (Fauske & Associates, LLC, 2019), and full spectrum LOCA (Westinghouse Electric Company, 2016).

Additional methodological advances to potentially explore in future work may include the linkage of GSA with state-of-the-art inverse problem methodologies such as developed by (Wang & Brigham, 2017) and (Walsh & Aquino, 2017). In principle, understanding of parameter sensitivities could prove beneficial in the context of challenging inverse problems, such as faced within nuclear reactor environments when many quantities of interest are not directly observable. Specifically, Figure 6.1 shows the configuration of a typical assembly of reactor internals highlighting some potential components and degradation mechanisms which could be explored via inverse mechanics methods per (Gerard, et al., n.d.) and (Westinghouse

Electric Company, 2009). For example, efficient quantification of the sensitivity imparted by monitor-able parameters at various locations (e.g., external vibration) on forms of degradation such as stress corrosion cracking of core barrel welds, or radial support wear could lead to enhanced methods of non-destructive evaluations or advanced sensing in the interest of structural health monitoring.

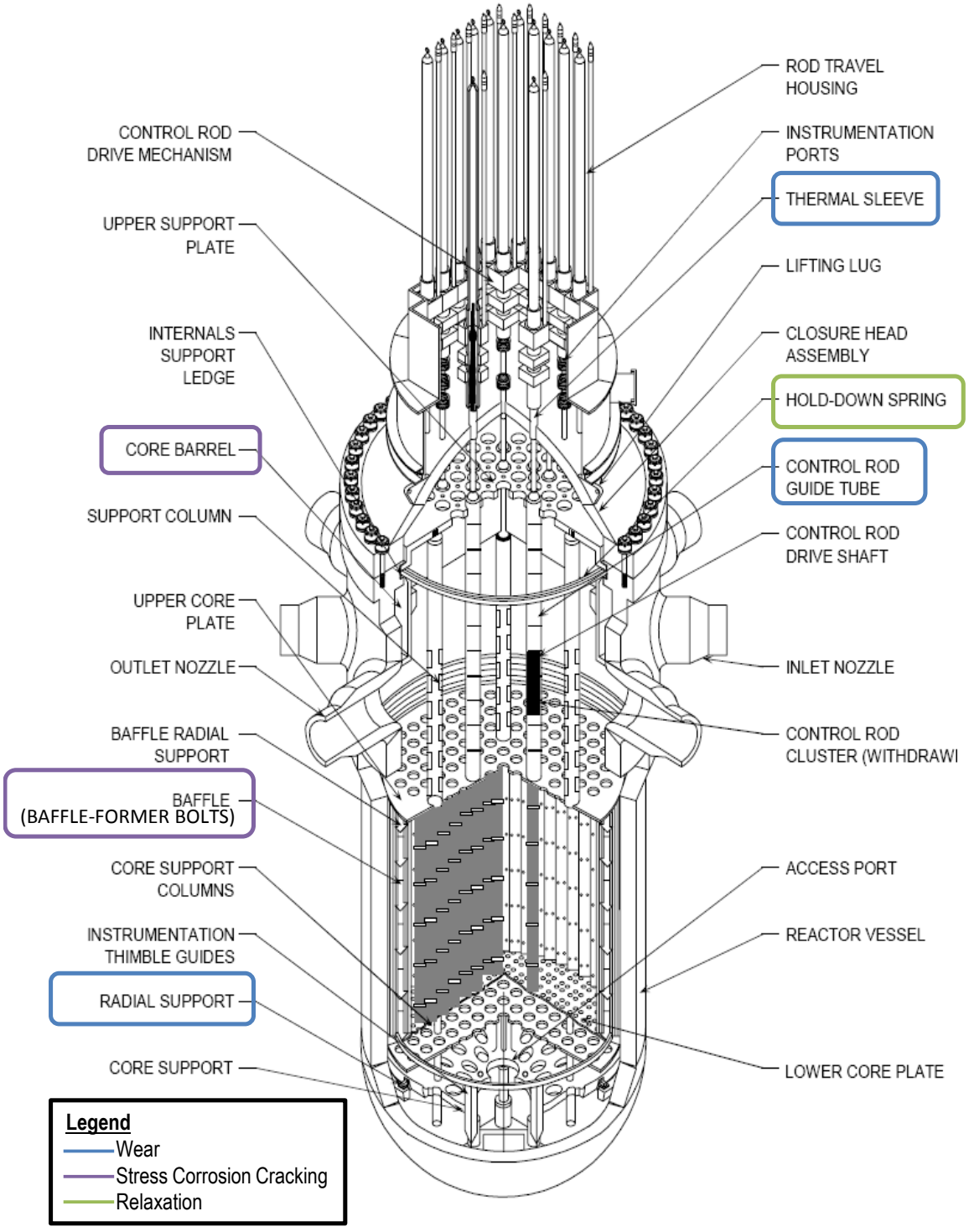

Figure 6.1 Illustration of Typical Reactor Internals Assembly with Potential Applications for Inverse Mechanics Problems involving Sensitivity Analysis

Appendix A

Matlab Code for Latinized Partially Stratified Sampling

In accordance with Section 2.4.2, this appendix shows the Matlab scripts used to accomplish LPSS.

The PSS design is first defined in terms of the sample design vectors. In both the UI FIV and SFEM analyses investigated within this dissertation, the PSS design included two sets of three subdomains. The strata were then defined with varying degrees of resolution within a switch statement. Upon defining the strata and subdomains, the LPSS function was called.

```
%%-------------------------------------------------------------------------%%
% Use of Partially Stratified Sampling (PSS) for Reactor Internals FIV    %
% PSS_UpperInternalsVibration.m - 09/13/2016                              %
% ... modified June 2017 to accomodate latinized PSS
% Author: Michael D. Shields, Civil Eng,JHU; modified by Greg Banyay      %
%                                                                         %
% This demonstration plots __ sample points of a 6 dimensional partially  %
% stratified sample design with three subdomains defined by pairing the   %
% direct-applied forcing function magnitudes & angles and the base motion %
% forcing function magnitudes & angles. The full 6-dimensional sample     %
% (given by variable x) is constructed by                                 %
% randomly pairing values from the three subdomains.                      %
%%-------------------------------------------------------------------------%%
clear;  clc;    close all;
% addpath(genpath('T:\Surrogate\LPSS'));
% addpath(genpath('T:\Surrogate\PSS'));
% addpath(genpath('T:\Surrogate\LSS'));

%% Partially stratified sample design vectors
pss_design =    [3,...      % cross-flow loads {x1 through x3}
                3];         % lateral & vertical base motion loads {x4, x5,
x6}

for ii = 1:7

    switch ii
        case 1; pss_strata =    [5, 5];
```

```
        case 2; pss_strata =    [8, 8];
        case 3; pss_strata =    [10, 10];
        case 4; pss_strata =    [12, 12];
        case 5; pss_strata =    [13, 13];
        case 6; pss_strata =    [3,  3];
        case 7; pss_strata =    [4,  4];% additional case for small DOE check
    end

    x{ii} = LPSS(pss_design,pss_strata);

    x_NotLatinized{ii} = PSS(pss_design,pss_strata);

    for i = 1:size(x{ii},2)

        figure(100+ii);
        subplot(6,1,i);         plot(x{ii}(:,i));  hold on;
                                plot(x_NotLatinized{ii}(:,i));
        legend('LPSS','PSS');   title(['parameter #', num2str(i)]);

    end

end
```

The LPSS function is defined as follows, within which the LSS function is called.

```
function [x lpss] = LPSS(lpss design,lpss strata)

%% Latinized Partially Stratified Sampling

nDim = sum(lpss_design);
nSamples = lpss_strata(1).^lpss_design(1);

col = 0;
for i = 1:length(lpss_design);
    nStrata = lpss_strata(i)*ones(1,lpss_design(i));
    x_lpss(:,col+1:col+lpss_design(i)) = LSS(nStrata);
    x_lpss(:,col+1:col+lpss_design(i)) =
_lpss(randperm(nSamples),col+1:col+lpss_design(i));
    col = col+lpss_design(i);
end
```

The LSS procedure is thus invoked with the following function.

```
function x = LSS(nStrata)

%% Copyright (C) Shields Uncertainty Research Group (SURG)
```

```
% All Rights Reserved
% Johns Hopkins University
% Department of Civil Engineering
% Updated: 16 July 2015
% Michael D. Shields

%% "Latinized" Stratified Sampling
% Generates a sample set on U(0,1) that is simultaneously a Latin hypercube
% sample and a stratified sample as described in:
% Shields, M.D. and Zhang, J. "The generalization of Latin hypercube
% sampling" Reliability Engineering and System Safety. (in review)

% Input:
% nStrata ----- Defines the design stratification
%               Vector (1 x nRVs)
%               Each column defines the number of stratifications to make
%                   in that direction.
% Example: nStrata = [25,25] produces a 2D design with 25x25
% stratification and 625 total samples
%
% Output:
% x ----------- Generated samples
%               Array (nSamples x nRVs)
%
%% Initialize variables
nSamples = prod(nStrata);
nRVs = length(nStrata);

%% Draw Latinized Stratified Samples

% Draw a Latin hypercube sample
x_temp = lhsdesign(nSamples,nRVs);
x = zeros(size(x_temp));

% Array to identify the candidate strata in each dimension for the LHS
% samples
strata = zeros(size(x_temp));

% Identify the candidate strata in each dimension for the LHS samples
for i = 1:nRVs
    y = 1/nStrata(i):1/nStrata(i):1;
    for j = 1:nSamples
        for k = 1:length(y)
            if x_temp(j,i) < y(k) && strata(j,i) == 0
                strata(j,i) = k;
                continue
            end
        end
    end
end

% Place the components of the LHS into the strata one-by-one
for i = 1:length(nStrata)
    l = 1;
    for j = 1:prod(nStrata(1:i))
        if l > nStrata(i)
            l = 1;
```

```
        end
        for k = 1:prod(nStrata)/prod(nStrata(1:i))
            for m = 1:length(strata)
                if strata(m,i) == l
                    x((j-1)*(prod(nStrata)/prod(nStrata(1:i)))+k,i) =
x_temp(m,i);
                    strata(m,i) = 0;
                    break
                end
            end
        end
        l = l+1;
    end
end
```

Appendix B

Discussion of Spectral Response Method

The method employed herein was itself verified against the method attributed to Jennings in (Nigam & Jennings, 1969). Jennings likewise uses a single degree-of-freedom oscillator model and thus starts with the same equation of motion as Equation (19), but uses β instead of α for the damping term and $a(t)$ instead of $\ddot{z}$ for the base acceleration term; his model is shown in Figure 6.2.

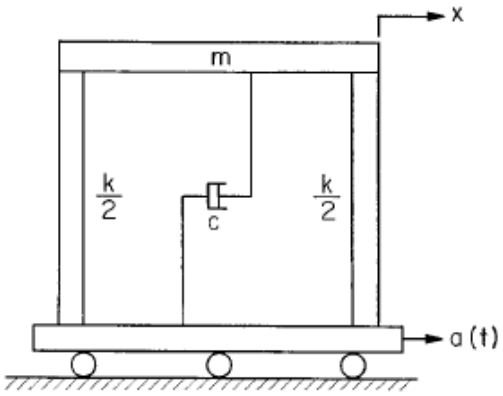

FIG. 1. A simple oscillator.

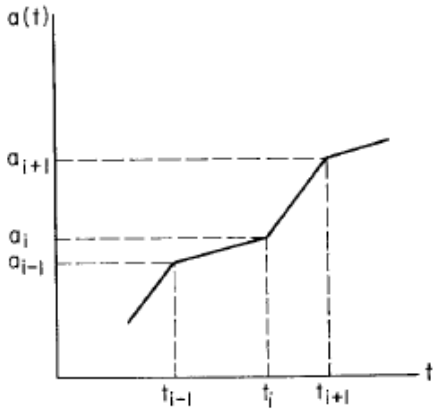

FIG. 2. Idealized base acceleration.

Figure 6.2 Jennings' Method Oscillator Model (Nigam & Jennings, 1969)

Jennings then concludes with an equation of equivalent form to Equation (25), and then goes on to express the acceleration response spectrum as:

$$S_a(\omega, \beta) = \max_{i=1,N}[\ddot{z}_i(\omega, \beta)] \tag{56}$$

Where $\ddot{z}_i$ is the acceleration of the mass at time t_i. Note that Jennings traces the use of response spectra from strong-motion earthquakes to Benioff, Biot, and Housner publications spanning from 1934 to 1941. Also, Jennings co-authored (Jennings, et al., 1968) along with Housner and Tsai in which they develop envelope functions for simulation of earthquakes and, in

so doing, attribute one mathematical model for use in generating "accelerograms" to a Shinozuka paper from 1967.

That approach is similar to that referred to as the Spectral Response Method (SRM) in numerous publications such as (Shinozuka & Deodatis, 1991), (Lang, et al., 2007), and (Shields, 2010). In (Shinozuka & Deodatis, 1991), the concept of the SRM is attributed to S.O. Rice in 1954 and state that Shinozuka and Jan first applied the SRM to simulation of a multi-dimensional, multi-variate nonstationary problem in 1972. The SRM is described as follows.

While (Shinozuka & Deodatis, 1991) did not necessarily pioneer the SRM, as they state, their paper serves as a detailed review of the method by gathering and compiling material that could only be found in several different papers. Starting with a one-dimensional univariate stationary stochastic process, $f_0(t)$ with autocorrelation function $R_{f_0 f_0}(\tau)$ and two-sided power spectral density function $S_{f_0 f_0}(\omega)$, (Shinozuka & Deodatis, 1991) provide a detailed derivation of the expression shown here as Equation (57), where Φ_k is an independent random phase angle uniformly distributed in the range $[0, 2\pi]$.

$$f_0(t) = \sqrt{2} \sum_{k=0}^{\infty} \left(2 S_{f_0 f_0}(\omega_k) \Delta\omega\right)^{\frac{1}{2}} \cos(\omega_k t + \Phi_k) \tag{57}$$

For simulation purposes, the i^{th} sample function may be expressed as:

$$f^{(i)}(t) = \sqrt{2} \sum_{n=0}^{N-1} \left(2 S_{f_0 f_0}(\omega_n) \Delta\omega\right)^{\frac{1}{2}} \cos\left(\omega_n t + \Phi_n^{(i)}\right) \tag{58}$$

The n^{th} frequency ω_n and frequency increment $\Delta\omega$ may be determined by the upper cut-off frequency of interest ω_u (i.e., that above which the PSD magnitude may be considered negligible) and the value of N chosen for the summation:

$$\omega_n = n\Delta\omega, \qquad n = 0,1,2, \dots, N-1 \tag{59}$$

$$\Delta\omega = {\omega_u}/{N} \tag{60}$$

The Fast Fourier Transform (FFT) may be used to accomplish this in a more computationally efficient manner

$$f^{(i)}(p\Delta t) = Re\left\{\sum_{n=0}^{M-1} \sqrt{2}\left(2S_{f_0f_0}(n\Delta\omega)\Delta\omega\right)^{\frac{1}{2}} e^{i\phi_n^{(i)}} e^{inp2\pi/M}\right\}, \quad p = 0,1,\dots,M-1 \tag{61}$$

To express this in a more concise form, let:

$$A_n = \left(2S_{f_0f_0}(n\Delta\omega)\Delta\omega\right)^{\frac{1}{2}}, \quad n = 0,1,\dots,M-1 \tag{62}$$

$$B_n = \sqrt{2}A_n e^{i\phi_n^{(i)}}, \quad n = 0,1,\dots,M-1 \tag{63}$$

Therefore, combining equations (61), (62), and (63) provides:

$$f^{(i)}(p\Delta t) = Re\left\{\sum_{n=0}^{M-1} B_n e^{inp2\pi/M}\right\}, \quad p = 0,1,\dots,M-1 \tag{64}$$

Note that in order to leverage the computational efficiency offered by the FFT, let $M = 2^{\mu}$ where μ is a positive integer.

In 2007, (Lang, et al., 2007) extended this work to address nonstationary stochastic processes. In so doing, they effectively re-derive Equation (57) in terms of a non-stationary PSD $S_{f_0f_0}(t,\omega_n)$ as:

$$f^{(i)}(t) = \sqrt{2}\sum_{n=0}^{N-1}\left(2S_{f_0f_0}(t,\omega_n)\Delta\omega\right)^{\frac{1}{2}}\cos\left(\omega_n t + \Phi_n^{(i)}\right) \tag{65}$$

Then, three different methods are offered by which a (non-stationary) evolutionary PSD may be formulated: the short-time Fourier transform, Wavelet transform, and Hilbert transform. The short-time Fourier transform (STFT) is chosen for use in this work. Note that the STFT method is consistent with the "spectrogram" function built into Matlab. The STFT may be written as:

$$S_{f_0 f_0}(t,\omega) = \iint_{-\infty}^{\infty} f^2(\tau) h^2(t-\tau)\, d\tau\, dt \tag{66}$$

where $h(t)$ is an appropriate window function (e.g., Hanning window) and τ is the time delay.

The evolutionary power spectrum (ES) is documented for the simulation of non-stationary stochastic processes by (Shields, 2010). Note that Section 3.2.1 of (Shields, 2010) cites (Jennings, et al., 1968) (written by the same P.C. Jennings as in (Nigam & Jennings, 1969)) as being one method by which the $A(t,\omega)$ (termed "envelope function" in (Jennings, et al., 1968)) can be evaluated. Assuming that the stationary PSD $S_{f_0 f_0}(\omega)$ exists such that $dS_{f_0 f_0}(\omega) = S_{f_0 f_0}(\omega) d\omega$ and a modulating function $A(t,\omega)$ is a real function of both frequency ω and time t, the ES may be written as (using the same nomenclature as the preceding discussion):

$$S_{f_0 f_0}(t,\omega) = A^2(t,\omega) S_{f_0 f_0}(\omega) \tag{67}$$

Given this discussion, it is understood that the SRM differs somewhat from the "accelerograms" referred to by (Jennings, et al., 1968) (i.e., if Shinozuka attributes the first use of the SRM to his work in 1972, then his prior work in 1967 was likely not equivalent to the SRM). Furthermore, the SRM is a method by which the spectrum, as a PSD, may be represented as a time-history. In conclusion, the SRM is effectively the means by which the time-history is developed for loading the SFEM, whereas the response spectrum method is used for plotting the resultant transient response on the frequency domain.

Matlab implementation of the USGS Open-File Report 92-296A (Converse & Brady, 1992) algorithm for computing acceleration response spectra is shown in the following table.

```
function RESP = ACCESLRS(arg1, arg2, arg3, arg4)
% arg1: acceleration time-history
% arg2: time step size for acceleration time-history
% arg3: vector of frequencies
% arg4: percent critical damping
p_tsteps =  length(arg1);
p_dt =      arg2;
p_fcount =  length(arg3);
p_d =       arg4;
for i = 1:p_fcount
    p_w =   2 * pi() * arg3(i);
    p_wd =  p_w * (1 - p_d^2)^0.5;
    p_w2 =  p_w^2;
    p_w3 =  p_w^3;

    p_zd =  0;
    p_zv =  0;
    p_za =  0;
    p_xd1 = 0;
    p_xv1 = 0;

    p_f1 =  (2.0 * p_d) / (p_w3 * p_dt);
    p_f2 =  1.0 / p_w2;
    p_f3 =  p_d * p_w;
    p_f4 =  1.0 / p_wd;
    p_f5 =  p_f3 * p_f4;
    p_f6 =  2.0 * p_f3;

    p_e =   exp(-p_f3 * p_dt);
    p_s =   sin(p_wd * p_dt);
    p_c =   cos(p_wd * p_dt);

    p_g1 =  p_e * p_s;
    p_g2 =  p_e * p_c;

    p_h1 =  p_wd*p_g2 - p_f3*p_g1;
    p_h2 =  p_wd*p_g1 + p_f3*p_g2;

    for ii = 1:(p_tsteps-1)
        p_dug =     arg1(ii+1) - arg1(ii);

        p_z1 =      p_f2 * p_dug;
        p_z2 =      p_f2 * arg1(ii);
        p_z3 =      p_f1 * p_dug;
        p_z4 =      p_z1 / p_dt;

        p_b =       p_xd1 + p_z2 - p_z3;
        p_a =       p_f4*p_xv1 + p_f5*p_b + p_f4*p_z4;

        p_xd2 =     p_a*p_g1 + p_b*p_g2 + p_z3 - p_z2 - p_z1;
        p_xv2 =     p_a*p_h1 - p_b*p_h2 - p_z4;

        p_xd1 =     p_xd2;
        p_xv1 =     p_xv2;

        p_aa =      -p_f6*p_xv1 - p_w2*p_xd1;
        p_f =       abs(p_xd1);
        p_g =       abs(p_xv1);
        p_h =       abs(p_aa);

        if p_f >= p_zd,     p_zd = p_f;     end
        if p_g >= p_zv,     p_zv = p_g;     end
        if p_h >= p_za,     p_za = p_h;     end

    end
    RESP(i,1) =     p_zd;
    RESP(i,2) =     p_zv;
    RESP(i,3) =     p_za;
end
end
```

Bibliography

Abbiati, G., Schobi, R., Sudret, B. & Stojadinivoc, B., 2017. *Structural Reliability Analysis using Deterministic Hybrid Simulations and Adaptive Kriging Metamodeling.* Santiago, Chile, Earthquake Engineering Research Institute.

Abrahamsen, P., 1997. *A Review of Gaussian Random Fields and Correlation Functions,* Oslo: Norwegian Computing Center.

Aistleitner, C., Hofer, M. & Tichy, R., 2012. A central limit theorem for Latin hypercube sampling with dependence and application to exotic basket option pricing. *International Journal of Theoretical & Applied Finance,* 15(7).

ANSYS, 2016. *Mechanical APDL Theory Reference.* 17.1 ed. Canonsburg(PA): SAS IP, Inc..

ASME, 2006. *Guide for Verification and Validation in Computational Solid Mechanics.* New York(NY): ASME.

ASME, 2009. *Standard for Verification and Validation in Computational Fluid Dynamics and Heat Transfer.* New York(NY): ASME.

ASME, 2017. *Rules for Construction of Nuclear Facility Components.* New York(NY): ASME.

Au-Yang, M. K., 2001. *Flow-Induced Vibration of Power and Process Plant Components.* New York(NY): ASME.

Au-Yang, M. K. & Connelly, W. H., 1977. A Computerized Method for Flow-Induced Random Vibration Analysis of Nuclear Reactor Internals. *Nuclear Engineering and Design,* Volume 42.

Au-Yang, M. K. & Jordan, K. B., 1980. Dynamic Pressure Inside a PWR – A Study Based on Laboratory and Field Test Data. *Nuclear Engineering and Design,* Volume 58, pp. 113-125.

Banyay, G. A., Ahmadpoor, M. & Brigham, J. C., 2014. *Proper Orthogonal Decomposition Based Reduced Order Modeling of the Very High Temperature Reactor Lower Plenum Hydrodynamics.* Chicago, IL, ASME.

Banyay, G. A., Brigham, J. C. & Rudnyi, E., 2012. *Investigation of Transient Thermal Analysis Computational Efficiency Improvements via Frequency Domain Methods.* Toronto, ON, ASME, pp. 861-869.

Banyay, G. A., Meyer, G. A. & Walker, A. P., 2015. *Proposed Changes to the ASME Boiler and Pressure Vessel Code Section III Appendix N for Flow-Induced Vibrations.* Boston, MA, ASME.

Banyay, G. A., Shields, M. D. & Brigham, J. C., 2019. Efficient Global Sensitivity Analysis for Flow-Induced Vibration of a Nuclear Reactor Assembly using Kriging Surrogates. *Nuclear Engineering and Design,* Volume 341, pp. 1-15.

Banyay, G. A., Smith, S. D. & Young, J. S., 2018. *Sensitivity Analysis of a Nuclear Reactor System Finite Element Model.* Minneapolis, MN, ASME.

Barker, B., 2017. *Nuclear Plant Life Extension: A Strategic Bridge.* [Online] Available at: http://eprijournal.com/nuclear-plant-life-extension-a-strategic-bridge/ [Accessed October 2018].

Ben Salem, M., Roustant, O., Gamboa, F. & Tomaso, L., 2017. Universal Prediction Distribution for Surrogate Models. *SIAM/ASA Journal on Uncertainty Quantification,* 5(1), pp. 1086-1109.

Ben Salem, M. & Tomaso, L., 2018. Automatic selection for general surrogate models. *Structural and Multidisciplinary Optimization.*

Bendat, J. S. & Piersol, A. G., 2010. *Random Data: Analysis and Measurement Procedures.* 4 ed. Hoboken(NJ): John Wiley & Sons.

Blevins, R. D., 2001. *Flow-Induced Vibration.* 2 ed. Malabar(FL): Kreiger.

Cannavo, F., 2012. Sensitivity analysis for volcanic source modeling quality assessment and model selection. *Computers & Geosciences,* Volume 44, pp. 52-59.

Chen, D., Levin, D. I., Sueda, S. & Matusik, W., 2015. Data-Driven Finite Elements for Geometry and Material Design. *ACM Transactions on Graphics,* August.34(4).

Cheng, K. et al., 2017. Global sensitivity analysis using support vector regression. *Applied Mathematical Modeling,* September.Volume 49.

Chrisman, L., 2014. *Latin Hypercube vs. Monte Carlo Sampling.* [Online] Available at: http://www.lumina.com/blog/latin-hypercube-vs.-monte-carlo-sampling#3 [Accessed December 2018].

Converse, A. & Brady, A. G., 1992. *BAP basic strong-motion accelerogram processing software version 1.0,* s.l.: United States Department of the Interior, U.S. Geological Survey.

Cukier, R. I., Levine, H. B. & Shuler, K. E., 1978. Nonlinear Sensitivity Analysis of Multiparameter Model Systems. *Journal of Computational Physics,* Volume 16, pp. 1-42.

Duchon, J., 1977. Splines minimizing rotation-invariant semi-norms in sobolev spaces. In: *Constructive theory of functions of several variables.* s.l.:Springer, pp. 85-100.

Fauske & Associates, LLC, 2019. *MAAP - Modular Accident Analysis Program.* [Online] Available at: https://www.fauske.com/nuclear/maap-modular-accident-analysis-program [Accessed January 2019].

Gerard, R., Bertolis, D. & Vissers, S., n.d. *Ageing Management of the reactor internals in Belgian nuclear units in view of Long Term Operation,* s.l.: International Atomic Energy Agency.

Gratiet, L. L., Marelli, S. & Sudret, B., 2016. *Metamodel-based Sensitivity Analysis: Polynomial Chaos Expansions and Gaussian Processes.* s.l.:ETH-Zurich.

Grigoriu, M., 2010. Linear random vibration by stochastic reduced-order models. *International Journal for Numerical Methods in Engineering,* Volume 82.

Grigoriu, M. & Field, R. V., 2014. A method for analysis of linear dynamic systems driven by stationary non-Gaussian noise with applications to turbulence-induced random vibration. *Applied Mathematical Modeling,* Volume 38.

Hattangady, N. V., 1999. Coarsening of Mesh Models for Representation of Rigid Objects in Finite Element Analysis. *International Journal for Numerical Methods in Engineering,* Volume 44, pp. 313-326.

Hossain, M. N., Sarkar, S. & Ghosh, D., 2015. Identification of dominant modes in random dynamical and aeroelastic systems. *Journal of Sound and Vibration,* Volume 357.

Hou, Z., Lu, W. & Chen, M., 2016. Surrogate-Based Sensitivity Analysis and Uncertainty Analysis for DNAPL-Contaminated Aquifer Remediation. *Journal of Water Resources Planning and Management,* November.142(11).

Huang, Z., Wang, C., Chen, J. & Tian, H., 2011. Optimal design of aeroengine turbine disc basd on kriging surrogate models. *Computers & Structures,* Volume 89.

Hughes, T. J., 1987. *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis.* Englewood Cliffs(NJ): Prentice Hall, Inc..

Idaho National Laboratory, 2015. *Reduced Order Model Implementation in the Risk-Informed Safety Margin Characterization Toolkit,* Idaho Falls: U.S. Department of Energy.

Iman, R. L., Davenport, J. M. & Zeigler, D. K., 1980. *Latin Hypercube Sampling (Program User's Guide),* Albuquerque, NM: Sandia Laboratories.

International Atomic Energy Agency, 2008. *Best Estimate Safety Analysis for Nuclear Power Plants: Uncertainty Evaluation.* [Online] Available at: https://www-pub.iaea.org/MTCD/publications/PDF/Pub1306_web.pdf [Accessed July 2018].

Jennings, P. C., Housner, G. W. & Tsai, N. C., 1968. *Simulated Earthquake Motions,* Pasadena: California Institute of Technology.

Jin, R., Chen, W. & Sudjianto, A., 2005. An Efficient Algorithm for Constructing Optimal Design of Computer Experiments. *Journal of Statistical Planning and Inference,* Volume 134.

Kaizer, J. S., 2013. *Fundamental Theory of Scientific Computer Simulation Review,* Rockville, MD: United States Nuclear Regulatory Commission.

Ko, D. Y. & Kim, K. H., 2012. Vibration and Stress Analysis of a UGT Assembly for the APR1400 RVI CVAP. *Nuclear Engineering and Technology,* Volume 44.

Ko, D. Y. & Kim, K. H., 2013. Structural Analysis of the CSB and LSS for APR1400 RVI CVAP. *Nuclear Engineering and Design,* Volume 261.

Lancaster, P. & Salkauskas, K., 1981. Surfaces Generated by Moving Least Squares Methods. *Mathematics of Computation,* July, 37(155), pp. 141-158.

Lang, J., Chaudhuri, S. R. & Shinozuka, M., 2007. Simulation of Nonstationary Stochastic Processes by Spectral Representation. *Journal of Engineering Mechanics,* June, 133(6), pp. 616-627.

Lataniotis, C., Marelli, S. & Sudret, B., 2017. *UQLab User Manual – Kriging (Gaussian Process Modelling).* Zurich: ETH-Zurich.

Loh, W.-L., 1995. *On the Convergence Rate to Normality of Latin Hypercube Sampling U-Statistics,* West Lafayette, IN: Purdue University.

Marelli, S. & Sudret, B., 2014. *UQLab: A framework for uncertainty quantification in Matlab.* Liverpool, UK, ASCE.

McKay, M. D., Conover, W. J. & Beckman, R. J., 1979. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics,* 21(2), pp. 239-245.

McRae, G. J., Tilden, J. W. & Seinfeld, J. H., 1980. Global Sensitivity Analysis – A Computational Implementation of the Fourier Amplitude Sensitivity Test (FAST). *Computers & Chemical Engineering.*

MIT Energy Initiative, 2018. *The Future of Nuclear Energy in a Carbon Constrained World.* Cambridge(MA): Massachusetts Institute of Technology.

Montgomery, D. C., 2013. *Design and Analysis of Experiments.* 8 ed. Hoboken(NJ): John Wiley & Sons.

Morgan, M. G., Henrion, M. & Small, M., 1992. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis.* Cambridge: Cambridge University Press.

Mulcahy, T. M., 1982. *Design Guide for Single Circular Cylinder in Turbulent Crossflow,* Lemont, IL: Argonne National Laboratory.

Nechak, L., Gillot, F., Besset, S. & Sinou, J. J., 2015. Sensitivity analysis and Kriging based models for robust stability analysis of brake systems. *Mechanics Research Communications,* Volume 69.

Nigam, N. C. & Jennings, P. C., 1969. Calculation of Response Spectra from Strong Motion Earthquake Records. *Bulletin of the Seismological Society of America*, April, 59(2).

Nuclear Energy Institute, 2014. *Aging Management Program Effectiveness.* Washington D.C.: Nuclear Energy Institute.

Oberkampf, W. L., Pilch, M. & Trucano, T. G., 2007. *Predictive Capability Maturity Model for Computational Modeling and Simulation,* Albuquerque, NM: Sandia National Laboratories.

Ortiz, K., Wirsching, P. H. & Paez, T. L., 1995. *Random Vibrations: Theory and Practice.* New York(NY): John Wiley & Sons.

Paez, T. L., Tucker, S. & O'Gorman, C., 1997. *Simulation of Nonlinear Random Vibrations using Artificial Neural Networks,* Albuquerque, NM: Sandia National Laboratory.

Palamara, M. J. et al., 2015. *Development of an Advanced PWR Reactor Internals System Finite Element Model for Flow-Induced Vibration Analyses.* Boston, MA, ASME.

Pianosi, F., Sarrazin, F. & Wagener, T., 2015. A Matlab toolbox for Global Sensitivity Analysis. *Environmental Modeling & Software,* Volume 70.

Rajakumar, C. & Rogers, C. R., 1991. The Lanczos Algorithm Applied to Unsymmetric Generalized Eigenvalue Problems. *International Journal for Numercial Methods in Engineering,* Volume 32, pp. 1009-1026.

Rasmussen, C. E. & Williams, C. K., 2005. *Gaussian Processes for Machine Learning.* Cambridge(MA): MIT Press.

Reddy, J. N., 2006. *An Introduction to the Finite Element Method.* 3 ed. New York(NY): McGraw Hill.

Sacks, J., Welch, W. J., Mitchell, T. J. & Wynn, H. P., 1989. Design and Analysis of Computer Experiments. *Statistical Science,* Volume 4, pp. 409-435.

Saltelli, A. et al., 2008. *Global Sensitivity Analysis: The Primer.* West Sussex, UK: John Wiley and Sons.

Schenk, C. A. & Schueller, G. I., 2005. *Uncertainty Assessment of Large Finite Element Systems.* Berlin: Springer-Verlag.

Shahsavani, D. & Grimvall, A., 2011. Variance-based sensitivity analysis of model outputs using surrogate models. *Environmental Modelling & Software,* June.26(6).

Shields, M. D., 2010. *Simulation of Stochastic Proceses: Applications in Civil Engineering,* Ann Arbor: ProQuest LLC.

Shields, M. D. & Zhang, J., 2016. The Generalization of Latin Hypercube Sampling. *Reliability Engineering & System Safety,* Volume 148, pp. 96-108.

Shinozuka, M. & Deodatis, G., 1991. Simulation of stochastic processes by spectral representation. *Applied Mechanics Reviews,* April, 44(4), pp. 191-203.

Smola, A. J. & Scholkopf, B., 1998. *A Tutorial on Support Vector Regression,* Berlin: NeuroCOLT2 Technical Report Series.

Sundar, V. S. & Shields, M. D., 2018. Reliability analysis using adaptive kriging surrogates with multimodel inference. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems: Part A Civil Engineering.*

Tennekes, H. & Lumley, J. L., 1972. *A First Course in Turbulence.* Cambridge(MA): MIT Press.

Tsai, N. C., 1972. Spectrum Compatible Motions for Design Purposes. *Journal of Engineering Mechanics,* April, 96(EM2), pp. 345-356.

United States Nuclear Regulatory Commission, 2013. *Programs for Monitoring Boiling-Water Reactor Steam Dryer Integrity.* [Online] Available at: https://www.nrc.gov/docs/ML1300/ML13003A049.pdf [Accessed July 2018].

United States Nuclear Regulatory Commission, 2014. *Regulatory Guide 1.60, Rev. 2.* Rockville(MD): United States Nuclear Regulatory Commission.

United States Nuclear Regulatory Commission, 2015. *Review of Lessons Learned from the San Onofre Steam Generator Tube Degradation Event.* [Online] Available at: https://www.nrc.gov/docs/ML1501/ML15015A419.pdf [Accessed July 2018].

United States Nuclear Regulatory Commission, 2017. *Comprehensive Vibration Assessment Program for Reactor Internals during Preoperational and Initial Startup Testing.* Rockville(MD): United States Nuclear Regulatory Commission.

Walsh, T. F. & Aquino, W., 2017. Massively parallel structural acoustics for forward and inverse problems. *Journal of the Acoustical Society of America,* Volume 141.

Wang, M. & Brigham, J. C., 2017. A generalized computationally efficient inverse characterization approach combining direct inversion solution initialization with gradient-based optimization. *Computational Mechanics,* Volume 59, pp. 507-521.

Westinghouse Electric Company, 1976. *Deblin 2, a Computer Code to Synthesize Earthquake Acceleration Time Histories (WCAP-8867),* Pittsburgh, PA: s.n.

Westinghouse Electric Company, 2001. *Nuclear Regulatory Commission.* [Online] Available at: https://www.nrc.gov/docs/ML0131/ML013190392.pdf [Accessed October 2018].

Westinghouse Electric Company, 2009. *PWR Vessel Internals Program Plan for Aging Management of Reactor Internals at Robinson Nuclear Plant (WCAP-17077-NP),* Pittsburgh, PA: Westinghouse Electric Company LLC.

Westinghouse Electric Company, 2011. *AP1000 Design Control Document.* Cranberry Township(PA): Westinghouse Electric Company, LLC.

Westinghouse Electric Company, 2015. *Comprehensive Vibration Assessment Program (CVAP) Vibration Analysis Program for the AP1000 Plant.* Cranberry Township(PA): Westinghouse Electric Company, LLC.

Westinghouse Electric Company, 2016. *Realistic LOCA Evaluation Methodology Applied to the Full Spectrum of Break Sizes (WCAP-16996-NP-A),* Pittsburgh, PA: Westinghouse Electric Company LLC.